University of Zurich

Department of Economics

Working Paper Series

ISSN 1664-7041 (print)
ISSN 1664-705X (online)

Working Paper No. 338

Quick or cheap? Breaking points in dynamic markets

Panayotis Mertikopoulos, Heinrich H. Nax and Bary S. R. Pradelski

December 2019

Quick or cheap?
Breaking points in dynamic markets*

Panayotis Mertikopoulos[◇], Heinrich H. Nax[§], and Bary S. R. Pradelski[‡]

December 30, 2019

Abstract

We examine two-sided markets where players arrive stochastically over time and are drawn from a continuum of types. The cost of matching a client and provider varies, so a social planner is faced with two contending objectives: *a*) to reduce players' *waiting time* before getting matched; and *b*) to form efficient pairs in order to reduce *matching costs*. We show that such markets are characterized by a *quick or cheap* dilemma: Under a large class of distributional assumptions, there is no 'free lunch', i.e., there exists no clearing schedule that is simultaneously optimal along both objectives. We further identify a unique breaking point signifying a stark reduction in matching cost contrasted by an increase in waiting time. Generalizing this model, we identify two regimes: one, where no free lunch exists; the other, where a window of opportunity opens to achieve a free lunch. Remarkably, greedy scheduling is never optimal in this setting.

Keywords: dynamic matching, online markets, market design

JEL classifications: D47, C78, C60, D80

*The paper has benefited from comments by Vahideh Manshadi, Igal Milchtaich, Jonathan Newton, Sven Seuken, Philipp Strack and seminar participants at Bar-Ilan University, the Paris Game Theory Seminar, and the INFORMS Workshop on Market Design 2019. We thank Simon Jantschgi and Dimitrios Moustakas for careful proof-reading. All errors are ours. PM benefited from the support of the COST Action CA16228 "European Network for Game Theory" (GAMENET). HN benefited from the support of the ERC Advanced Investigator Grant 'Momentum' (No. 324247). BP benefited from the support of the ANR grants 15-IDEX-02, 11LABX0025-01, ALIAS and the Oxford-Man Institute of Quantitative Finance.

[◇]Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, LIG, 38000 Grenoble, France; panayotis.mertikopoulos@imag.fr

[§]Univ. Zurich, 8050 Zurich, Switzerland; heinrich.nax@uzh.ch

[‡]Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, LIG, 38000 Grenoble, France; bary.pradelski@cnrs.fr

1 Introduction

Many economic interactions require the dynamic matching of heterogeneous agents that arrive stochastically to a two-sided market. Examples include the dynamic matching of clients and providers in markets for jobs and services, of buyers and sellers in financial markets, of taxis and passengers on road networks, of donors and recipients in organ exchanges, etc.[1]

It is known that many of these markets vary substantially in terms of efficiency (Roth and Xing, 1994, 1997). The focus of our investigation is on a crucial aspect of market design in this context, namely the *scheduling of clearing events*. The goal of "making a thick market" (Roth, 2012) is to find the best schedule of market clearing so that sufficient clients and providers are in the market to allow for stable and efficient matches over time while not waiting excessively. Designing an optimal clearing policy thus requires optimizing along the following two objectives:

1. To reduce the coexistence of agents on the two sides of the market.
2. To match parties in such a way so as to minimize cost (or maximize productivity).

In pursuit of these two goals, clearing schedules need to be formulated to address the following key question: *How long should the social planner wait between two clearing events?*

To illustrate the above, consider the example of a governmental employment bureau faced with a dynamically evolving job market where job offers and job seekers arrive to the system stochastically over time. The bureau has two aims, namely to reduce the coexistence of vacancies and job seekers, and to match vacancies with the skills of individual job seekers so as to maximize productivity. Waiting times incur costs via unemployment benefits, as well as costs due to productivity losses incurred by badly staffed vacancies.

To gain in generality, we abstract away from application-specific details (such as the particular structure of the application and recruitment processes). This allows us to focus on the trade-offs between two different and concurrent objectives, waiting time and matching cost. Perhaps surprisingly, this *quick or cheap* dilemma is not easily resolvable as greedy scheduling policies are generally not optimal in this context.

[1] More generally, we focus on markets for 'nondurables'; for classic studies on product durability and market performance see Smith (1962), Smith et al. (1988), and Dickhaut et al. (2012).

Related work

Dating back to the 1950s, the first related strand of work focuses on behavioral aspects underlying the dynamics of unemployment and job vacancies in labor markets (Dow and Dicks-Mireaux, 1958). These analyses identify avenues to reduce *waiting* – i.e., the coexistence of unemployment and vacancies – by better understanding the behavior of job seekers and job providers. Lines of reasoning proposed to explain the coexistence of unemployment and vacancies include the classical search models of McCall (1970), Mortensen (1970), and Lucas and Prescott (1974), as well as more recent models with workforce inertia due to Shimer (2007).[2] We complement this literature with a view that some degree of waiting is actually beneficial from a social welfare perspective as it enables market thickening – which in turn enables mismatch reduction. To illustrate this, consider the example of Shimer (2007), where some laid-off steel workers are not immediately given vacant positions as nurses. This may indeed be deemed optimal by a social planner when – by delaying their match – these nurse vacancies eventually are taken up by better nurses and the jobless steel workers find other jobs in the steel industry that might become available in the future.[3]

The second strand of related work comes from the matching literature and extends the canonical static matching framework to a dynamic setting.[4] As in the example of steel workers and nurses above, *mismatch* in dynamic environments may occur due to temporal inconsistencies, whereby, a posteriori, better matches were precluded by inferior matches that were formed earlier on. Therefore, some delay may be optimal from a social planner perspective in order to reduce mismatch. From a practical viewpoint, the challenge is to identify optimal mechanisms that thicken and clear the market in a way that balances these two objectives.

In this regard, Akbarpour et al. (2017), Ashlagi et al. (2019), Baccara et al. (2018), and Loertscher et al. (2018) break new ground in identifying optimal clearing schedules.[5] More precisely, Akbarpour et al. (2017), in the spirit of an organ ex-

[2]Note that Shimer (2007) terms his explanandum “mismatch” (as opposed to “waiting”), a term the matching literature uses to describe suboptimal matchings, which may be confusing.

[3] *Waiting* is explained behaviorally through inertia in Shimer (2007), that is, by the argument that steel workers stay close to their factories hoping that they reopen; Lucas and Prescott (1974) propose a different interpretation whereby waiting is due to the fact that steel workers must actively spend some time searching for these nursing jobs elsewhere.

[4]The canonical static frameworks underlying our analyses were pioneered by Egervary (1931); Koenig (1931), and Edmonds (1965); see also Gale and Shapley (1962) for matching with ordinal preferences.

[5]These are inspired by some earlier papers on dynamic matching in organ exchange by Uenver (2010); Zenios (2002). See Akbarpour et al. (2017) for a discussion. See also Bloch and Houy

change application such as the 'kidney exchange', identify the optimal mechanism to maximize the number of matches, that is, to minimize the number of agents perishing resulting from failing to get recipients matched with donors in time. In the model of Akbarpour et al. (2017), agents from both sides of the market arrive and leave stochastically and all carry identical match values, i.e., they are of the same type (in the spirit of each life being worth the same). However, there are two types of agents, as some matches are feasible and others are infeasible, thus rendering some agents easier – others harder – to match.[6] The optimal mechanism identified by Akbarpour et al. (2017) minimizes the number of unmatched patients based on information concerning arrivals and departures, which may involve delaying compatible matches. Without such information, greedy scheduling is always optimal in this setting. In fact, Ashlagi et al. (2019) show that greedy policies are generally optimal, even if information about departure times is available when the above kind of 'kidney exchange' markets becomes large.

In a related setting, Baccara et al. (2018) and Loertscher et al. (2018) introduce binary diversifications of agents and the notion of waiting costs instead of perishing rates as in Akbarpour et al. (2017). In Baccara et al. (2018), agents arrive in donor-recipient pairs and recipients are allowed to decline matches in order to remain in the market. This is motivated by the applications under scrutiny which include, among others, child adoption. As a result, one of the study's key focuses is on strategic incentives and their role in determining market outcomes. Their optimal clearing policy is discriminatory, in that it involves matching same-type pairs greedily, and delaying up to some threshold when there are only cross-type pairs in the market. By contrast, Loertscher et al. (2018) introduce a common discount factor (instead of a constant waiting cost) for both types of agents (as well as for the social planner).[7] They focus on the analysis of the possibility of efficient trade and rent extraction by the market maker.

In theoretical computer science, the study of related questions dates back at least to the pioneering paper of Karp et al. (1990).[8] To the best of our knowledge,

(2012), Kurino (2014), and Leshno (2012) who study related queuing models where one side of the market is already present (such as in the housing market).

[6]This can be modeled by means of a dynamically changing compatibility graph where edges represent feasible matches.

[7]The discount factor is motivated by the study's focus on financial markets, and would be determined by risk-free rate, beta, and risk premium. Related analyses of financial markets include Budish et al. (2015) and Wah et al. (2015) (see also Wah et al. (2017) for market-making more generally) who consider periodic clearing of order books to abate certain market phenomena such as volatility that stem from high-frequency trading.

[8]Karp et al. (1990) and subsequent work – similar to its economic counterparts – focus on

Emek et al. (2016) were the first in this strand of research to consider the scenario where all agents arrive on the market over time (instead of just one market side). They present a non-bipartite model where requests arrive stochastically from one of n different locations to study the performance of different algorithms in terms of worst-case matching and waiting cost.[9] In the setting of Emek et al. (2016), the specific match costs result from the distance between agents' locations so that, for a patient social planner, it is optimal to wait and only match agents who are at the same location.

Finally, motivated by ride-sharing applications, Ashlagi et al. (2017) extend the model of Emek et al. (2016) to a bipartite setting where agents independently arrive at different locations. Ashlagi et al. (2017) study the performance of a family of clearing schedules along the two axes of waiting vs. mismatch separately, an approach that we extend in order to formulate the induced trade-off between waiting time and the cost of matching.

Contributions of the paper

Our paper examines dynamic markets with an infinite type space (in contrast to one or two types), a framework we call the *dynamic clearing game*. Our point of departure is the static assignment game of Shapley and Shubik (1972) to which we add a dynamic layer whereby clients and providers arrive to the market stochastically and independently. Skills are drawn from a large class of type distributions so that every match is possible but some matches are more costly than others. At each matching event, the social planner must decide who to match with whom, and how long to wait before the next matching event. As such, the social planner is called to weigh, on the one hand, mismatches incurred from matching clients and providers suboptimally; and, on the other hand, the agents' waiting time. To address this dual issue, we study clearing schedules in terms of *when* to match in a fully heterogeneous setting where the social planner has no information regarding the cost of matching couples currently in the market and has no information about the future arrivals of individuals.

---

models with two market sides, where by contrast one side is typically present to begin with and incoming agents from the other side can only match with some of the present agents according to a compatibility graph (see Mehta (2013) for an overview and Aggarwal et al. (2011) for extensions to vertex-weighted matching).

[9] Azar et al. (2017) obtain additional results in terms of upper and lower bounds for the original model. Emek et al. (2019) obtain sharper results for a two-location model. There are also other extensions such as allowing for a stochastic graph (Anderson et al., 2015; Ashlagi et al., 2018).

For concreteness, we start by studying a micro-level model where costs of individual matches are distributed according to independent exponential random variables. Whilst our results hold for other distributions too, this model has the advantage of being tractable in closed form. In more detail, we first establish a class of optimal clearing schedules for two extreme types of single-objective social planners – that is, for social planners who only care about minimizing waiting time (in which case greedy is best) or mismatch costs (resulting in endless delay), but not both at the same time. Second, we show that these two objectives are mutually incompatible, and multi-objective social planners (who care about both) face a fundamental trade-off. Specifically, there is no 'free lunch', that is *there is no clearing schedule that is approximately optimal in terms of both waiting time and matching cost.* Remarkably, the greedy clearing schedule is sub-optimal for every multi-objective social planner.

Building on the no free lunch result, we proceed to fill the spectrum between matching cost and waiting time minimization. We do so by introducing a class of clearing schedules covering a wide range of social planning desiderata between waiting time and matching cost, and achieving a continuous trade-off between the two. To explore the finer aspects of this trade-off, we introduce a utility model for the social planner whereby the associated utility of matching cost is of the same order as the agents' utility of waiting time. Under this model, we show that there exists a non-trivial clearing schedule achieving this balance, and we show that this schedule is effectively unique (up to asymptotic order considerations).

Finally, we generalize our key findings by studying different decay rates of matching costs (instead of focusing on one decay rate that results from the micro-founded match costs). We identify two regimes. One, where no free lunch continues to hold. The other, where the benefit from waiting is growing quickly enough, such that a window of opportunity opens and it *is* possible to get a free lunch. As before, in both regimes, greedy scheduling is generally sub-optimal.

Compared to the existing literature on the trade-off between waiting and mismatch (both in economics and computer science) our model introduces *incomplete information* about the distribution of past and future match costs and considers an infinite type space in a tractable model. As a consequence, the social planner tries to resolve the trade-off between matching optimally and waiting time in light of incomplete information. Incomplete information in our setting implies that the social planner must employ clearing schedules that do not take as input the relative strengths of current and future matches (since the latter is unknown), thus yielding qualitatively new results.

In contrast to prior results for markets with one or two types of match costs where lack of information resulted in optimality of some form of greedy scheduling, we find that greedy clearing is generally not optimal in the presence of many types. Hence, the quick-versus-cheap trade-off is more intricate than previously found. Moreover, our results may actually also have consequences for applications that have been studied before too (e.g. kidney exchange) if other match value metrics (e.g. potential years of life lost or disability-adjusted life years) are used that would produce more than binary match values. By studying fully heterogeneous match costs we have to rely on different mathematical tools compared to previous analyses, which were often able to reduce the induced dynamics to discrete Markov processes.

The key technical innovations of our paper concern the concurrent consideration of a continuum of types, independent arrivals, and incomplete information. In turn, these contributions rely on a range of previously unused tools from probability theory and disordered systems to obtain closed-form solutions. These underlying results are concerned with the expected matching cost for given instances of random, static assignment games. In particular, in static assignment games with the same number of clients and providers and exp(1) distributed edge weights, Aldous (2001) proved the long-standing conjecture that the expected minimum weight matching converges to $\pi^2/6$ (i.e., as the number of players is growing). This result was later extended by Waestlund (2005) to assignment games with match costs drawn from non-identical exponential distributions.[10] By leveraging the techniques of Aldous (2001) and Waestlund (2005), we are able to compute the expected matching cost for every 'snapshot in time' of the dynamic clearing game. This provides strong foundations for our proofs which are then focused on estimating the fluctuations that result from the random arrival of clients and providers and their randomly drawn match costs. To achieve this, we use several approximation techniques (in particular, the approximation of the arrival process by a continuous-time Wiener process), which allow us to port over several results from martingale limit theory (such as the law of the iterated logarithm).

[10] To the best of our knowledge, the work of Walkup (1979) is the first to pose the question, while Mezard and Parisi (1987) conjectured the specific limit value. We also leverage the analyses of Buck et al. (2002) and Linusson and Waestlund (2004) who obtain results for the expected values of finite instances of the latter models, showing – as a byproduct – that the value is increasing with the number of agents. For a survey of this literature, we refer the reader to Krokhmal and Pardalos (2009).

Paper outline. The rest of the paper is structured as follows. In Section 2, we introduce the *dynamic clearing game* and the performance measures relevant for our analysis. In Section 3, we show that a social planner who cares about both waiting and mismatch faces a fundamental and non-negligible trade-off. We then go on to analyze a natural selection of clearing schedules in Section 4, which cover the whole range of possible trade-offs. In Section 5, we commit to a specific utility function that specifies how the social planner values waiting time versus matching cost and find the unique optimal clearing schedule. Section 6 generalizes the analysis and shows that there are two regimes, one where 'free lunch' is not achievable and one where it is achievable. Finally, in Section 7, we discuss practical implications and of avenues for future research.

2 The model

In this section, we introduce the model, which we shall refer to as the *dynamic clearing game*.

The dynamic clearing game. Consider the following model of a dynamic two-sided market evolving in continuous time $\tau \in [0, \infty)$. At each tick of a Poisson clock with rate 1 an agent enters the market; this agent could be either a *client* or a *provider*, with equal probability.[11] To keep track of the number of agents in both sides of the market, let $\mathcal{C}(\tau)$ and $\mathcal{P}(\tau)$ denote the set of clients and providers that have entered the market by time τ (and possibly already left again), and let $N_{\mathcal{C}}(\tau) = |\mathcal{C}(\tau)|$ and $N_{\mathcal{P}}(\tau) = |\mathcal{P}(\tau)|$ be the respective numbers thereof. Then, the number of agents on the *short side of the market* will be written $N(\tau) = \min\{N_{\mathcal{C}}(\tau), N_{\mathcal{P}}(\tau)\}$.[12]

As in the static assignment model of Shapley and Shubik (1972) on which we build, we consider a one-to-one matching market where each client is to be *matched* to at most one provider and vice versa; then, once a couple is matched, both agents leave the market. For example, in labor market language, each job seeker gets

[11] We are using here the terms 'client' and 'provider' in a generic sense, just to illustrate the difference between the two sides of the market. As we explain below, what is important from a modeling perspective is that 'clients' are to be matched to 'providers' (as in our running example of job vacancies and job seekers).

[12] In a slight (but convenient) abuse of notation, we will sometimes write $N_{\mathcal{C}}(t)$, $N_{\mathcal{P}}(t)$, and $N(t)$ to denote respectively the number of clients, providers, and agents at the short side of the market when the t-th agent enters the market – specifically, letting $\tau(t)$ denote the time at which the t-th agent enters the market, we will write $N_{\mathcal{C}}(t) \equiv N_{\mathcal{C}}(\tau(t))$, etc.

at most one job, and each vacancy concerns exactly one worker; once a match has been made, the governmental job bureau removes the matched pair from its ledger, and the process continues.

For concreteness, we shall next define a specific family of match cost distributions. This is done solely to streamline our presentation: Our methodology allows us to be more general as we discuss in Section 6. Suppose that the quality of a (candidate) pair is characterized by an inherent *match parameter* λ_{ij} where a higher parameter will represent a lower expected match cost. Match costs are independently and exponentially distributed with rate λ_{ij}.[13]

Specifically, we posit that the *match cost* $w_{ij} > 0$ when client $i \in \mathcal{C}$ is matched to provider $j \in \mathcal{P}$ is an independent draw from an exponential distribution of rate λ_{ij} for any time τ, that is, $w_{ij} \sim \exp(\lambda_{ij})$. For example, a popular model assumes that λ_{ij} is composed by additively separable components describing the agents' types and a couple-specific term depending possibly on both the identity of the agents and their types (Kanoria et al., 2018). For generality, our only assumption regarding the rate parameters λ_{ij} is that they are bounded from below by $\underline{\lambda}$, from above by $\overline{\lambda}$, and have mean value λ, that is, $\lambda = \lim_{\tau\to\infty}[N_{\mathcal{C}}(\tau) + N_{\mathcal{P}}(\tau)]^{-1} \sum_{i=1}^{N_{\mathcal{C}}(\tau)} \sum_{j=1}^{N_{\mathcal{P}}(\tau)} \lambda_{ij}$.

The social planner. Throughout the sequel, we assume the existence of a social planner who, whenever an agent arrives on the market, observes the arrival; other than that, the social planner has no other information regarding the arrival process of the agents (or the distribution of their match costs). Due to this lack of information, the social planner has no basis to judge whether a particular agent arriving in the market is 'good' or 'bad', and is thus left with the challenge of choosing a clearing schedule with which to operate the market. In the sequel, we will also write $A \equiv A(\tau)$ for the number of clients/providers that have been assigned a partner up to time τ, and $R(\tau) = N_{\mathcal{C}}(\tau) + N_{\mathcal{P}}(\tau) - 2A(\tau)$ for the number of unmatched agents up to time τ.

With all this in hand, a *clearing schedule* (CS) will be a rule that determines:

(i) At which points in time $\tau \in (0, \infty)$ to trigger a *matching event* (ME), possibly depending on $N_{\mathcal{C}}(\tau)$, $N_{\mathcal{P}}(\tau)$ and $A(\tau)$.

[13] As mentioned by Aldous (2001) and developed in detail by Janson (1999, Section 2) generalizations to larger classes of distributions are easily obtained. For ease of exposition we stick to exponential distributions with the exception of Section 6 that generalizes our main results.

(ii) Which players to match at a given matching event, possibly depending on the current match costs of agents who have already arrived to the market until time τ.

After a matching event, the players who are being matched leave the market, while the unmatched players remain on the market.

In what follows, we shall focus on clearing schedules that match a single couple per matching event. In particular, the clearing schedules we analyze will match the couple with the minimal matching cost in each matching event.[14] Restricting ourselves to these kinds of clearing schedules is motivated by our aim to study clearing schedules that can be paired with a market mechanism (e.g., a two-sided auction). Myerson and Satterthwaite (1983) and Rustichini et al. (1994) show the general impossibility to have ex-post efficient and budget balanced mechanisms for two-sided market games with private information. Their results rely on the assumption that, with positive probability, any given client-provider pair have valuations for each other such that trade is not individually rational for both at any price. Since the latter doesn't hold in our setting, one could micro-found the interaction avoiding the impossibility, that is, define a mechanism that is ex-post efficient and budget balanced. We shall assume throughout the analysis the existence of such a mechanism; however, given our focus on the social planner an explicit analysis of such mechanisms is beyond the scope of the present paper.

Now, as discussed before, the social planner aims to match clients and providers optimally along two axes: *a*) to reduce the coexistence of clients and providers (i.e., *waiting time*); and *b*) to match clients and providers in a way that minimizes matching cost (i.e., *mismatch*). Beginning with the latter, the *expected matching cost* for the first A couples is defined as

$$\mathrm{cost}_{\mathsf{CS}}(A) \equiv \mathbb{E}\left[\sum_{k=1}^{A} w_{i_k,j_k}\right] \tag{1}$$

where w_{i_k,j_k} is the match cost of the k-th matched couple and the expectation is taken with respect to the random arrival of clients and providers and the randomness of the match costs. Similarly, the *expected waiting time* of a clearing schedule

[14] The only clearing schedule that we consider and which violates this principle is the first-come, first-served (FCFS) clearing schedule which we describe in Section 3.

until time T is defined as

$$\text{wait}_{\texttt{CS}}(T) \equiv \mathbb{E}\left[\int_0^T R(\tau)\, d\tau\right] \tag{2}$$

where the expectation is taken with respect to the random arrival of agents in the market. In the sections that follow, we will explore the optimization of these two performance metrics, and the trade-offs that arise when trying to minimize both.

3 The trade-off between waiting time and matching costs

Our analysis begins with the case of a single-minded social planner. Specifically, we investigate which clearing schedule a social planner would employ if either only caring about the expected waiting time, or only caring about the expected matching cost. After we deal with these two cases separately, we shall proceed to show that these objectives are mutually incompatible and lead to an unavoidable trade-off for the social planner.

Single-minded social planners. First, a social planner who is optimizing the agents' expected waiting time will choose a clearing schedule which leaves no unmatched couples at any point in time. To do so, we will consider a 'greedy' clearing schedule, denoted $\texttt{CS}_{\texttt{greedy}}$, which performs a minimum weight matching whenever there is exactly one unmatched client/provider pair in the market. Second, a social planner who is optimizing the agents' expected matching cost will choose a clearing schedule which – ideally – waits until everyone has arrived in the market and then matches agents optimally (thus minimizing the sum of match costs).[15] That is, the hypothetical 'patient' clearing schedule, denoted $\texttt{CS}_{\texttt{patient}}$, should be preferred by any social planner who is only concerned with the expected matching cost.

The implementation of these schedules leads to the following matching cost and waiting time:

[15] To make such a clearing schedule realistic, all agents would need to arrive in the market in finite time; since this schedule will mostly serve as a theoretical comparison baseline, we will not consider this issue in detail.

Proposition 1. *The optimal clearing schedules for a single-objective social planner are:*

(1) *The patient clearing schedule* $\mathtt{CS}_{\mathtt{patient}}$ *is optimal with respect to matching cost minimization; in particular, for all $A \geq 1$, we have:*

$$\frac{\log 2}{\overline{\lambda}} \leq \mathrm{cost}_{\mathtt{patient}}(A) \leq \frac{\pi^2}{6\underline{\lambda}}$$

(2) *The greedy clearing schedule* $\mathtt{CS}_{\mathtt{greedy}}$ *is optimal with respect to waiting time minimization; in particular, for all $\tau \geq 0$, we have:*

$$\mathrm{wait}_{\mathtt{greedy}}(\tau) = \frac{2}{3}\tau^{3/2}$$

Remark. In view of Proposition 1, the expected matching cost of $\mathtt{CS}_{\mathtt{patient}}$ and the expected waiting time of $\mathtt{CS}_{\mathtt{greedy}}$ will serve as the benchmark for comparing the matching cost and waiting time of any other clearing schedule.

Proof of Proposition 1. We prove our claims for each of the two clearing schedules separately.

Part 1: Matching cost minimization. For the first assertion, note that the exponential distribution is closed under scaling by a positive factor, i.e., if $X \sim \exp(\kappa)$ then $\mu X \sim \exp(\kappa/\mu)$. In our case, this implies that

$$w_{ij} \sim \exp(\lambda_{ij}) \iff w_{ij} \sim \frac{1}{\lambda_{ij}} \exp(1) \tag{3}$$

We have that for all $i \in \mathcal{C}$, $j \in \mathcal{P}$, the distribution of w_{ij} is first-order stochastically dominated by $\underline{\lambda}^{-1} \exp(1)$. Thus the expected weight is upper bounded by the simplified problem where all match costs are distributed according to $\underline{\lambda}^{-1} \exp(1)$. With this in mind, we will simplify notation in the rest of the proof by setting $\underline{\lambda} = 1$.

By the summation formula of Buck et al. (2002) and Linusson and Waestlund (2004), we have for the expected weight of the minimum A-matching (note that $A = N$, recalling that $N = \min\{N_{\mathcal{C}}, N_{\mathcal{P}}\}$):

$$\mathbb{E}_{\min}\left[\sum_{k=1}^{N} w_{i_k,j_k}\right] = \sum_{\substack{i,j \geq 0 \\ i+j<N}} \frac{1}{(N_{\mathcal{C}} - i) \cdot (N_{\mathcal{P}} - j)}. \tag{4}$$

Thus, we readily have

$$\mathrm{cost}_{\texttt{patient}}(N) = \mathbb{E}_{\min}\left[\sum_{k=1}^{N} w_{i_k,j_k}\right] \leq \sum_{\substack{i,j\geq 0\\ i+j<N}} \frac{1}{(N-i)\cdot(N-j)}. \tag{5}$$

To proceed, by Waestlund (2009, Lemma 3.1) we have

$$\sum_{\substack{i,j\geq 0\\ i+j<N}} \frac{1}{(N-i)\cdot(N-j)} = \sum_{k=1}^{N} \frac{1}{k^2} \leq \zeta(2), \tag{6}$$

where $\zeta(2) = \sum_{n=1}^{\infty} 1/n^2 = \pi^2/6$ is the Basel constant. Returning to our original problem, we conclude that $\mathrm{cost}_{\texttt{patient}}(A) \leq \pi^2/(6\underline{\lambda})$, as claimed.

To compute the lower bound, we need to consider the expected match cost for $A = 1$, because matching more players would only serve to increase the expected matching cost. In the patient clearing schedule for $A = 1$, the process terminates when at least one client and at least one provider have entered the market. Let $Y \geq 2$ be the number of agents required to observe at least one client and one provider. Then the event $Y = k+1$ is the same event as the union of the disjoint events 'the first k agents are clients and the $(k+1)$-th agent is a provider' and 'the first k agents are providers and the $(k+1)$-th agent is a client'. Each of the latter events has probability $1/2^{k+1}$, so $\mathbb{P}(Y = k+1) = 2^{-k}$. Moreover, as we prove in Lemma 8, for $Y = k+1$, the expected minimum match cost is given by $\frac{1}{\bar{\lambda}\cdot k}$. Thus for $A = 1$, we get

$$\mathrm{cost}_{\texttt{patient}}(N) = \sum_{k=1}^{\infty} \frac{1}{\bar{\lambda}k}\,\mathbb{P}(Y = k+1) = \frac{1}{\bar{\lambda}} \sum_{k=1}^{\infty} \frac{1}{2^k k} = \frac{\log 2}{\bar{\lambda}}, \tag{7}$$

where the last equality follows from the series expansion $\log(1-x) = -x - x^2/2 - x^3/3 - \cdots$ applied to $x = 1/2$.

Part 2: Waiting time minimization. For our second assertion, note that, at any point in time, there are either no clients or no providers in the market. In view of this, let S_τ be the difference of clients and providers who have arrived to the market until time τ, that is, $S_\tau = N_{\mathcal{C}}(\tau) - N_{\mathcal{P}}(\tau)$. Then, for all $T > 0$, we get:

$$\mathrm{wait}_{\texttt{greedy}}(T) = \mathbb{E}\left[\int_0^T |S_\tau|\, d\tau\right] = \int_0^T \mathbb{E}[|S_\tau|]\, d\tau \tag{8}$$

where the latter equality holds by Tonelli's theorem (since $|S_\tau|$ is non-negative).

Applying Tonelli's theorem a second time, we can consider the case where the expectation with respect to the arrival times is taken first. To do so, consider the process where at the fixed points in time $\tau = 1, 2, \dots$ an agent arrives to the market and let $\bar{S}_\tau$ be the difference of clients and providers who have arrived to the market at time τ. We then have:

$$\mathbb{E}\left[\int_0^T |S_\tau|\, d\tau\right] = \int_0^T \mathbb{E}[|S_\tau|]\, d\tau = \int_0^T \mathbb{E}[|\bar{S}_\tau|]\, d\tau \tag{9}$$

It is well-known that for $\tau \to \infty$ the appropriately rescaled random walk $\bar{S}_\tau$ converges in distribution to the Wiener process W_τ (Kac, 1947). Thus, for large T, Eq. (9) gives

$$\mathbb{E}\left[\int_0^T |S_\tau| d\tau\right] = \int_0^T \mathbb{E}[|W_\tau|]\, d\tau = \int_0^T \sqrt{\mathrm{Var}(W_\tau)}\, d\tau = \int_0^T \sqrt{\tau}\, d\tau = \frac{2}{3}T^{3/2}. \quad \square$$

This concludes the analysis of a single-minded social planner who either only cares about the expected waiting time, or onlythe expected matching cost.

Multi-objective social planners. Going beyond the narrow view of a single-minded social planner, we proceed below to examine the case of social planners that care about *both* the expected cost of matching and the agents' overall expected waiting time. Natural candidates to evaluate a clearing schedule in this context are the expected matching ratio and the expected waiting ratio, defined below as follows:

1. The *expected matching ratio* of a clearing schedule $\texttt{CS}$ is

$$\alpha \equiv \alpha(A) = \frac{\mathrm{cost}_{\texttt{CS}}(A)}{\mathrm{cost}_{\texttt{patient}}(A)}. \tag{10}$$

2. The *expected waiting ratio* of a clearing schedule $\texttt{CS}$ is

$$\beta \equiv \beta(\tau) = \frac{\mathrm{wait}_{\texttt{CS}}(\tau)}{\mathrm{wait}_{\texttt{greedy}}(\tau)}. \tag{11}$$

Note that $\texttt{CS}_{\texttt{patient}}$ is optimal with respect to expected matching cost while $\texttt{CS}_{\texttt{greedy}}$ is optimal with respect to expected waiting time.

Going forward, note that all candidate clearing schedules can be characterized by a function $f\colon \mathbb{R}^+ \to \mathbb{R}^+$ such that the k-th ($k \in \mathbb{N}$) couple is matched when $\lceil f(k) \rceil$ agents are on the short side of the market. Denote a clearing schedule that is defined via a function f by $\mathtt{CS}_f$. Without any restrictions on f this includes all possible clearing schedules that always match the couple with the minimal match cost. However, given that we wish to analyze the asymptotic regime where enough agents have entered the market, we focus below on a natural class of functions introduced by Hardy (1910) which make such comparisons possible. Specifically, each function in this class is defined, for all $x \geq 0$, by a finite combination of the basic arithmetic operations (addition, multiplication, raising to a power, and their inverses), operating on the variable x and on real constants. Hardy (1910, Theorem, page 18) shows that for any two such functions, f and g, either $f = \omega(g), f = \Theta(g)$, or $f = o(g)$.

Then, for such functions, we will use the following asymptotic notations:
$f(x) = \mathcal{O}(g(x))$ if $f(x) < c \cdot g(x)$ for some $c > 0$ constant and x sufficiently large.
$f(x) = \Omega(g(x))$ is the inverse O notation ($f(x) > c \cdot g(x)$ for x sufficiently large).
$f(x) = \Theta(g(x))$ if there exist two constants $k, K \geq 0$ and a positive integer x_0 such that $kg(x) \leq f(x) \leq Kg(x)$ for all $x \geq x_0$.
For $g(x)$ non-zero $f(x) = o(g(x))$ if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$ and $f(x) = \omega(g(x))$ if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = \infty$.

In light of the above, the first question that arises is whether there exists a clearing schedule that is optimal along *both* axes (at least, asymptotically). To formalize this, we say that a clearing schedule $\mathtt{CS}$ has *finite expected matching ratio* if $\limsup_{A(\tau)>0} \alpha(A(\tau)) < \infty$; likewise, we say that clearing schedule has *finite expected waiting ratio* if $\limsup_{\tau>0} \beta(\tau) < \infty$.

The following theorem shows that the answer to the above question is a resounding 'no':

Theorem 2 ('No free lunch'). *There exists no clearing schedule simultaneously achieving finite ratios for both expected matching cost and waiting time.*

Theorem 2 illustrates that multi-objective social planners are faced with a crucial trade-off independently of their specific utility function – provided of course that they care about both matching cost and waiting time in a non-trivial way. In addition, Theorem 2 justifies the performance measures for the two dimensions of mismatch (expected matching ratio and waiting ratio) and in particular the sufficiency to analyze them in terms of orders of τ (or $A = A(\tau)$).

Proof of Theorem 2. We prove this result by contradiction; specifically, we find a necessary condition for a clearing schedule to have finite expected matching ratio and then show that clearing schedules satisfying this condition cannot have a finite expected waiting ratio.

To make this precise, consider the clearing schedule $\texttt{CS}_f$ that matches the k-th couple when $N - (k-1) = \lceil f(k) \rceil$. We shall show that a necessary condition for a clearing schedule $\texttt{CS}_f$ to have finite expected matching ratio is

$$f(k) = \omega(k^{1/2}) \tag{12}$$

To show this, consider the clearing schedule that matches the k-th couple when at least $\lceil k^{1/2} \rceil$ players are on the short side of the market; with a fair degree of hindsight, denote this clearing schedule as $\texttt{CS}_{\gamma=1/2}$.[16] As we show in Theorem 3, this clearing schedule has $\alpha(A) = \Theta(\log A)$. Thus, in order for another schedule $\texttt{CS}_f$ to have finite expected matching ratio, matching events have to happen orders of magnitude later than in $\texttt{CS}_{\gamma=1/2}$. Concretely, for k large enough the k-th couple is cleared at time $\tau_{\texttt{CS}_f}(k) = \omega(\tau_{\texttt{CS}_{\gamma=1/2}}(k))$. It follows that Eq. (12) holds.

We can now analyze the expected waiting time for $\texttt{CS}_f$ such that Eq. (12) holds for f. To construct a lower bound, consider the alternative arrival process, where clients and providers alternatingly arrive to the market. Note that for any given clearing schedule this process incurs lower waiting time. For the clearing schedule we consider the waiting time of this alternative arrival process is precisely governed by the fact that the k-th match takes place when at least $f(k)$ players are on the short side of the market. Further, note that $\tau(A)$ is clearly smaller for this new arrival process compared to the original process. Given that we only need to show that the waiting time is increasing in A it suffices to show that it is increasing in τ (not conditioned on A). Thus, the waiting time is lower bounded by using the approximation by the Wiener process (as in the proof of Proposition 1) and by observing that arrival is governed by a Poisson clock of rate 1:

$$\int_0^T 2f(\tau)d\tau = \omega\left(\int_0^T 2\sqrt{\tau}\, d\tau\right) = \omega(T^{3/2}) \tag{13}$$

By Proposition 1(ii), the optimal expected waiting time is $(2/3)\,T^{3/2}$, so we conclude that the expected waiting ratio is lower bounded by $\omega(T^{3/2})/T^{3/2} = \omega(1)$ and our proof is complete. □

[16] See Section 4 for detailed definitions.

This concludes our first result for multi-objective social planners, showing that the trade-off between cost of matching and waiting time is essential.

4 Interpolating between waiting time and matching cost

In this section, we analyze a class of clearing schedules covering a broad spectrum of social planning desiderata interpolating between matching cost and waiting time.

To begin, recall that Proposition 1 provides the expected matching cost of the patient clearing schedule $\mathtt{CS}_{\mathtt{patient}}$ (which minimizes mismatches) and the expected waiting time of the greedy clearing schedule $\mathtt{CS}_{\mathtt{greedy}}$ (which minimizes waiting times). Interpolating between these two 'extreme' schedules, we shall consider below a class of clearing schedules where the social planner waits for some length of time in order to accrue some intermediate number of agents on both sides of the market. Concretely, we shall study clearing schedules that match the k-th couple when $N - k = f(k)$, i.e., when $f(k)$ players are on the short side of the market.[17]

For concreteness, we restrict ourselves to clearing schedules of the form

$$f(k) = \Theta(k^\gamma) \quad \text{for some } \gamma \in [0,1]. \tag{14}$$

For $\gamma = 0$, the induced clearing schedules match players once a constant threshold is reached; in particular, the greedy schedule is recovered when $f \equiv 1$ (corresponding to $\gamma = 0$). More generally, we shall denote clearing schedules of the above form by $\mathtt{CS}_\gamma$ and write $\mathtt{CS}_{\gamma=1/2}$ for the clearing schedule with $\gamma = 1/2$. Similarly we shall use the notation α_γ for the expected matching ratio of $\mathtt{CS}_\gamma$ and β_γ for the expected waiting ratio of $\mathtt{CS}_\gamma$.

In addition to the clearing schedules induced by the assumptions above, we shall also consider another natural schedule based on the principle of *first-come, first-served* (FCFS), i.e., when agents are matched as soon as possible on a first-come, first-served basis. This schedule, which we denote by $\mathtt{CS}_{\mathtt{FCFS}}$, differs from $\mathtt{CS}_{\mathtt{greedy}}$ in terms of who is matched with whom (first-come, first-served vs. minimum cost matching) but not regarding when a matching event occurs. As such, given that

[17] Recall that $N = \min\{N_\mathcal{C}, N_\mathcal{P}\}$.

$\texttt{CS}_{\texttt{FCFS}}$ does not take into account matching costs, it is not reasonable to expect that it will perform well on any dimension other than the agents' expected waiting times. On the other hand, it exhibits 'fairness' relative to the agents' arrival times, a feature which is crucial in many applications.[18]

Overview of results. Table 1 summarizes all clearing schedules analyzed below (including a 'balanced' schedule, $\texttt{CS}_{\texttt{balanced}}$, that we discuss in Section 5). Our results (in terms of each schedule's expected matching and waiting ratio) are then summarized in Table 2: as can be seen, the family of schedules under study captures the full range between schedules that are 'good' relative to mismatches and 'bad' relative to waiting times, and vice versa.

$\texttt{CS}_{\texttt{FCFS}}$	match players as soon as possible on a first-come, first-served basis
$\texttt{CS}_{\texttt{greedy}}$	match players as soon as possible
$\texttt{CS}_{\gamma}$	match the k-th couple when $\Theta(k^\gamma)$ players are on the short side of the market $(0 \leq \gamma \leq 1)$
$\texttt{CS}_{\texttt{patient}}$	match players optimally after everyone has arrived
$\texttt{CS}_{\texttt{balanced}}$	match the k-th couple when $\Theta(k^{1/2}(\log k)^{1/3})$ players are on the short side of the market.

Table 1: Overview of the various clearing schedules considered in the sequel. Note that all schedules other than $\texttt{CS}_{\texttt{FCFS}}$ match the couple with the minimum match cost at each matching event.

In view of these results, the clearing schedule $\texttt{CS}_{\gamma=1/2}$ can be seen as a *phase transition* between two markedly different regimes. On the one hand, for $\gamma < 1/2$, the expected matching ratio $\alpha(A)$ grows as a power law in A while the expected waiting ratio $\beta(\tau)$ is finite. On the other hand, for $\gamma > 1/2$, we have a finite expected matching ratio but an expected waiting ratio that grows polynomially. Finally, at the critical point $\gamma = 1/2$, the expected matching ratio grows to infinity for large A, but at a slow, logarithmic rate ($\Theta(\log A)$). Notably, the phase transition at $\gamma = 1/2$ signifies a discontinuity of the expected matching ratio, so it is a *first-order* phase transition; by contrast, the expected waiting ratio exhibits no such discontinuity, signifying a *second-order* phase transition.

The infinite matching ratio vis-a-vis the finite waiting ratio for $\gamma = 1/2$ suggests

[18] Indeed, this may be a desirable feature in applications such as processor time requests in distributed computing. We shall leave extensions of our analyses to include fairness considerations for future work.

Schedule	Description	Matching ratio, α	Waiting ratio, β
$\texttt{CS}_{\texttt{FCFS}}$	FCFS matching	$\Theta(A)$	1
$\texttt{CS}_{\texttt{greedy}}$	Greedy matching	$\Omega(A^{1/2})$	1
$\texttt{CS}_{0\leq\gamma<1/2}$	Subcritical rate matching	$\Omega(A^{1/2-\gamma})$	$\Theta(1)$
$\texttt{CS}_{\gamma=1/2}$	Critical rate matching	$\Theta(\log A)$	$\Theta(1)$
$\texttt{CS}_{1/2<\gamma\leq 1}$	Supercritical rate matching	$\Theta(1)$	$\Theta(\tau^{\gamma-1/2})$
$\texttt{CS}_{\texttt{patient}}$	Patient matching	1	$\Theta(\tau^{1/2})$
$\texttt{CS}_{\texttt{balanced}}$	Balanced matching	$\Theta((\log A)^{1/3})$	$\Theta((\log A)^{1/3})$

Table 2: The range of expected matching and waiting ratios; $\texttt{CS}_{\texttt{balanced}}$ is discussed in Section 5. Recall from Eq. (10) and Eq. (11) the expected matching ratio $\alpha \equiv \alpha(A) = \frac{\mathrm{cost}_{\texttt{CS}}(A)}{\mathrm{cost}_{\texttt{patient}}(A)}$ ($\mathrm{cost}_{\texttt{patient}}(A) = \Theta(1)$) and the expected waiting ratio $\beta \equiv \beta(\tau) = \frac{\mathrm{wait}_{\texttt{CS}}(\tau)}{\mathrm{wait}_{\texttt{greedy}}(\tau)}$ ($\mathrm{wait}_{\texttt{greedy}}(\tau) = \Theta(\tau^{3/2})$).

that further fine-tuning should be possible and, indeed, the 'balanced' schedule $\texttt{CS}_{\texttt{balanced}}$ (which we define and discuss in Section 5) reduces the growth of the expected matching ratio by a factor of $(\log A)^{2/3}$ while increasing the expected waiting ratio $\beta(\tau)$ by a factor of $(\log \tau)^{1/3}$. In a sense (that we shall make precise in the next section) this is as close as we can get to a 'free lunch' in this setting.

Formal statements. We now proceed to provide complete statements of the results discussed above. To streamline our presentation, we have relegated the detailed proofs to Appendices A and B; however, the main pattern of the proofs can also be seen in Section 5 where we treat the case of $\texttt{CS}_{\texttt{balanced}}$.

We begin with our results for the matching cost ratio α:

Theorem 3. *The expected matching ratios for the schedules under study are as follows:*

($\mathtt{CS}_{\mathtt{FCFS}}$) *FCFS matching:*
$$\alpha_{\mathtt{FCFS}} = \frac{6\underline{\lambda}\lambda}{\pi^2} A \tag{15a}$$

($\mathtt{CS}_{\mathtt{greedy}}$) *Greedy matching:*
$$\alpha_{\mathtt{greedy}} \geq \frac{6\underline{\lambda}}{5\pi^2} A^{1/2} \tag{15b}$$

($\mathtt{CS}_{0\leq\gamma<1/2}$) *Subcritical rate matching:*
$$\underline{C}_\gamma\, A^{1/2-\gamma} \leq \alpha_{0\leq\gamma<1/2} \leq \overline{C}_\gamma\, A^{1-2\gamma} \tag{15c}$$

($\mathtt{CS}_{\gamma=1/2}$) *Critical rate matching:*
$$\frac{2\underline{\lambda}}{\pi^2} \log A \leq \alpha_{\gamma=1/2} \leq \frac{\overline{\lambda}}{\underline{\lambda}} \frac{1+\log A}{\log 2} \tag{15d}$$

($\mathtt{CS}_{1/2<\gamma\leq 1}$) *Supercritical rate matching:*
$$\alpha_{1/2<\gamma\leq 1} = \frac{\overline{\lambda}}{\underline{\lambda}} \frac{\zeta(2\gamma)}{\log 2} \tag{15e}$$

($\mathtt{CS}_{\mathtt{patient}}$) *Patient matching:*
$$\alpha_{\mathtt{patient}} = 1 \tag{15f}$$

Remark. In the above, $\underline{C}_\gamma$ and $\overline{C}_\gamma$ are positive constants, and $\zeta(s) = \sum_{n=1}^{\infty} 1/n^s$ denotes the Riemann zeta function (so $\zeta(s) < \infty$ for all $s > 1$).

By contrast, for the expected waiting ratio β, we have:

Theorem 4. *The expected waiting ratios for the schedules under study are as follows:*

($\mathtt{CS}_{\mathtt{FCFS}}$) *FCFS matching:*
$$\beta_{\mathtt{FCFS}} = 1 \tag{16a}$$

($\mathtt{CS}_{\mathtt{greedy}}$) *Greedy matching:*
$$\beta_{\mathtt{greedy}} = 1 \tag{16b}$$

($\mathtt{CS}_{0\leq\gamma<1/2}$) *Subcritical rate matching:*
$$\beta_{0\leq\gamma<1/2} = \Theta(1) \tag{16c}$$

($\mathtt{CS}_{\gamma=1/2}$) *Critical rate matching:*
$$\beta_{\gamma=1/2} = \Theta(1) \tag{16d}$$

($\mathtt{CS}_{1/2<\gamma\leq 1}$) *Supercritical rate matching:*
$$\beta_{1/2<\gamma\leq 1} = \Theta(\tau^{\gamma-1/2}) \tag{16e}$$

($\mathtt{CS}_{\mathtt{patient}}$) *Patient matching:*
$$\beta_{\mathtt{patient}} = \Theta(\tau^{1/2}) \tag{16f}$$

In closing this section, it is worth noting that the bounds for α become asymptotically 'less tight' for small $\gamma < \frac{1}{2}$. As far as this gap is concerned, we conjecture that the upper bound is the tight one: the lower bound is obtained via a crude approximation using Jensen's inequality, and this could be potentially tightened (although we haven't been able to do so). By contrast, the approximation for the upper bound seems less drastic.

We should also note that the results in Theorem 4 are driven by the assumption that the arrival of either a client or a provider at every stage of the process is equally likely. This entails that the expected absolute difference of clients and providers $|N_{\mathcal{C}}(\tau) - N_{\mathcal{P}}(\tau)|$ can by approximated by a Wiener process as detailed in Appendix B. For the latter we know that the expectation is $\sqrt{\tau}$, so $|N_{\mathcal{C}}(\tau) - N_{\mathcal{P}}(\tau)| \approx \sqrt{\tau}$ in expectation. It would be interesting to consider different arrival processes such as an urn model with delayed replacement where $|N_{\mathcal{C}}(\tau) - N_{\mathcal{P}}(\tau)|$ would exhibit a different asymptotic behavior; we leave this analysis to future work.

5 A balanced social planner

Until now, we have analyzed clearing schedules based on the trade-off between waiting time and matching cost, but without explicitly comparing the two. In this section, we shall commit to a specific class of utility functions in order to make an explicit comparison between these otherwise incomparable quantities.

To that end, let $u(\cdot)$ denote the expected utility (or 'welfare') of the social planner given a specific clearing schedules. Assume further that the functions expressing this utility depend on both the expected matching cost and the expected waiting time via the additively separable expression

$$u(\texttt{CS}) = u_{\text{cost}}(\alpha_{\texttt{CS}}) + u_{\text{wait}}(\beta_{\texttt{CS}}) \tag{17}$$

In order to make comparisons between the utility components u_{cost} and u_{wait} we shall first consider their respective maximum values. It is then natural to assume that u_{cost} is maximal for the patient clearing schedule (which minimizes matching cost) and that u_{wait} is maximal for the greedy clearing schedule (which minimizes waiting time). We shall thus assume that the two maxima are of the same order, viz.,

$$\sigma \cdot u_{\text{cost}}(\alpha_{\texttt{patient}}) = (1-\sigma) \cdot u_{\text{wait}}(\beta_{\texttt{greedy}}) \tag{18}$$

where $\sigma \in (0,1)$ is a constant factor that specifies the relative importances of the disutilities from mismatching versus waiting. Naturally, we require that the social planner seeks to minimize both the costs of matching and the agents' waiting time. As such, we make the assumption that u_{cost} is a concave function that decreases in the expected matching cost, and u_{wait} is a concave function that decreases in the expected waiting time.

In view of all this, a social planner is said to be *balanced* if the disutilities from mismatching and waiting display similar growths for large τ. That is, for a given clearing schedule $\texttt{CS}$ with expected matching ratio α and expected waiting time β, we assume that

$$u_{\text{cost}}(\alpha_{\texttt{CS}}) = \Theta(u_{\text{wait}}(\beta_{\texttt{CS}})) \quad \text{whenever} \quad \alpha_{\texttt{CS}} = \Theta(\beta_{\texttt{CS}}) \tag{19}$$

In this general context, we obtain the following result governing balanced social planning:

Theorem 5. *Let* $\texttt{CS}_{\texttt{balanced}}$ *be the clearing schedule that matches the* k*-th couple when* $N-(k-1) = \lceil k^{1/2}(\log k)^{1/3}\rceil$ *players are on the short side of the market. The expected matching and waiting ratios incurred by* $\texttt{CS}_{\texttt{balanced}}$ *are both* $\Theta((\log A)^{1/3})$*; moreover, any other schedule* $\texttt{CS}_f$ *achieving this balance has* $f(k) = \Theta(k^{1/2}(\log k)^{1/3})$.

Remark. For technical reasons we state our result in terms of the number of matched couples ($A = A(\tau)$). Note that, for any clearing schedule where the proportion of matched players increases over time (more precisely, where $\lim_{\tau\to\infty} \frac{A(\tau)}{N_{\mathcal{C}}(\tau)+N_{\mathcal{P}}(\tau)} = 1$), A is growing at the same rate as τ.

Proof of Theorem 5. Consider a clearing schedule of the form $\texttt{CS}_f$ that matches the k-th couple when $\lceil f(k)\rceil$ players on the short side of the market. In order to balance the expected matching and waiting ratios, any such clearing schedule would have to satisfy $f(k) = \omega(\sqrt{k})$; otherwise, the expected matching ratio would dominate asymptotically the expected waiting ratio (see Table 2). Thus, without loss of generality, we can assume that $f(k)$ is non-decreasing for large k.

Let $t(k, f(k))$ be the stopping time for the event that for the k-th time at least $f(k)$ clients and $f(k)$ providers are in the market, assuming that every time this is the case one client and one provider are removed. Finally, recall that $S_\tau = N_{\mathcal{C}}(\tau) - N_{\mathcal{P}}(\tau)$ is the difference of clients and providers who have arrived to the market until τ.

We begin with the expected matching ratio. For the upper bound we have:

$$\begin{aligned}\mathbb{E}\left[\sum_{k=1}^{A}\frac{1}{(\lceil f(k)\rceil+|S_{t(k,f(k))}|)\lceil f(k)\rceil}\right] &= \sum_{k=1}^{A}\mathbb{E}\left[\frac{1}{(\lceil f(k)\rceil+|S_{t(k,f(k))}|)\lceil f(k)\rceil}\right]\\ &\leq \sum_{k=1}^{A}\lceil f(k)\rceil^{-2}\end{aligned} \tag{20}$$

For the lower bound, an algebraic argument which we make precise in Appendix A (cf. Eq. (58)) shows that $\mathbb{E}[t(k,f(k))] < 10k$. Furthermore, note that $\mathbb{E}[S_t]$ is strictly increasing in t. Thus, by Jensen's inequality, and Lemma 9 (which is bounding $|S_t|$ via a combinatorial argument and using Stirling's formula), we get:

$$\begin{aligned}\sum_{k=1}^{A}\mathbb{E}\left[\frac{1}{(\lceil f(k)\rceil+|S_{t(k,f(k))}|)\lceil f(k)\rceil}\right] &\geq \sum_{k=1}^{A}\frac{1}{(\lceil f(k)\rceil+\mathbb{E}[|S_{10k}|])\lceil f(k)\rceil}\\ &\geq \frac{\pi}{\sqrt{2}e}\sum_{k=1}^{A}\frac{1}{(\lceil f(k)\rceil+\sqrt{10k})\lceil f(k)\rceil}\\ &= \Theta\left(\sum_{k=1}^{A}\frac{1}{f(k)^2}\right)\end{aligned} \tag{21}$$

where the last line follows from the assumption $f(k)=\omega(k^{1/2})$. Thus the two bounds together with the fact that the patient schedule has finite matching cost yield the result that, for $f(k)=\omega(k^{1/2})$ the expected matching ratio is $\alpha(A) = \Theta\left(\sum_{k=1}^{A} 1/[f(k)^2]\right)$.

We proceed, by considering the incurred waiting time. Recall that $N(\tau)-A(\tau)$ is the number of agents on the shorter side of the market at time τ, so $N-A = \lceil f(k-1)\rceil - 1$ after the $(k-1)$-st match. The number of clients and the number of providers that need to arrive to the market before the k-th match is thus upper bounded by

$$\lceil f(k)\rceil - \big(\lceil f(k-1)\rceil - 1\big) = 1 + \lceil f(k)\rceil - \lceil f(k-1)\rceil \leq 2 \tag{22}$$

where the last inequality follows from the fact that $f(k)=o(k)$ is a necessary condition for a feasible clearing schedule (that is, a clearing schedule where the proportion of unmatched versus matched players is decreasing). The expected waiting time accrued between the $(k-1)$-st and the k-th match is therefore upper bounded by:

$$\Delta_k := \underbrace{\mathbb{E}[\text{time s.t. } \geq 2 \text{ clients } \& \geq 2 \text{ providers enter market}]}_{=:\Delta_k^1} \cdot \underbrace{\big(2\lceil f(k) \rceil + \lceil g(k) \rceil\big)}_{=:\Delta_k^2} \tag{23}$$

where $\lceil g(k) \rceil$ is a function that we will use to upper bound $|S_k|$, viz., the random variable constituting the absolute difference of clients and providers in the market at time $\tau(k)$. For posterity, note also that Δ_k^1 is the expectation of the time between the $(k-1)$-th and the k-th match and Δ_k^2 provides an upper bound for the number of agents waiting in the time interval between the $(k-1)$-th and the k-th match.

Given the arrival of agents is governed by a Poisson clock of rate one, we have $\Delta_k^1 = 5$, i.e., on average, five agents need to enter the market to have at least two clients and at least two providers. To see this, let Y be the number of flips of a coin required to observe at least 2 heads (clients) and 2 tails (providers). The event '$Y > k$' is then equivalent to the union of the events '$\binom{k}{k-1}$ heads' and '$\binom{k}{k-1}$ tails'. The two latter events are disjoint and each has probability $\frac{k}{2^k}$. Thus $\mathbb{P}[Y > k] = \frac{k}{2^{k-1}}$ and we have

$$\begin{aligned} \mathbb{E}[Y] &= \sum_{k=0}^{\infty} \mathbb{P}(Y > k) = 1 + 2\sum_{k=1}^{\infty} \frac{k}{2^k} = 1 + 2\sum_{k=1}^{\infty}\sum_{j=1}^{k} \frac{1}{2^k} && (24) \\ &= 1 + 2\sum_{j=1}^{\infty}\sum_{k=j}^{\infty} \frac{1}{2^k} = 1 + 2\sum_{j=1}^{\infty} \frac{1}{2^{j-1}} = 5 && (25) \end{aligned}$$

We thus have for Eq. (23)

$$\Delta_k = 5 \cdot \big(2\lceil f(k) \rceil + \lceil g(k) \rceil\big) \tag{26}$$

Next, to choose the function $g(k)$, note that the law of the iterated logarithm gives

$$\lim_{k \to \infty} \frac{|S_k|}{\sqrt{2k \log \log k}} = 1. \tag{27}$$

Hence, by choosing $g(k) = \sqrt{2k \log \log k}$, the random variable $|S_k|$ is asymptotically bounded from above by $g(k)$ with probability one.

We consider two cases below, which are exhaustive by Hardy (1910, Theorem, page 18):

Case 1: $f(k) = \Omega(g(k))$. For the first case we have:

$$5 \cdot (2f(k) + g(k)) = \Theta(f(k)) \tag{28}$$

The expected waiting ratio until A pairs have been matched is bounded from above by $A^{-3/2} \sum_{k=1}^{A} \Theta(f(k))$, where we are using the fact that the expected waiting time for the greedy schedule is given by $A^{3/2}$ (see Proposition 1). A trivial lower bound for the expected waiting ratio is then given by

$$\frac{1}{A^{3/2}} \sum_{k=1}^{A} 2f(k) = \frac{1}{A^{3/2}} \sum_{k=1}^{A} \Theta(f(k)) \tag{29}$$

Thus, the expected waiting ratio is given by

$$\beta(A) = \Theta\left(\frac{1}{A^{3/2}} \sum_{k=1}^{A} f(k) \right) \tag{30}$$

Moving to the comparison of matching and waiting ratios, we recall that u_{cost} and u_{wait} are decreasing and concave and are of the same order (by assumption). Thus $u = u_{\text{cost}} + u_{\text{wait}}$ is maximized if and only if $\alpha = \Theta(\beta)$. In turn, this holds if and only if

$$\sum_{k=1}^{A} \frac{1}{f(k)^2} = \Theta\left(\frac{1}{A^{3/2}} \sum_{k=1}^{A} f(k) \right) \tag{31}$$

or, equivalently, if and only if

$$\int_1^A \frac{1}{f(x)^2}\, dx = \Theta\left(\frac{1}{A^{3/2}} \int_1^A f(x)\, dx \right) \tag{32}$$

where the asymptotic passage from summation to integration – i.e., from Eq. (31) to Eq. (32) – is made precise in Appendix C.

We shall show that $f(x) = \Theta(\sqrt{x}(\log x)^{1/3})$ is the unique solution to Eq. (32) up to order. To simplify notation, let $f(x) = \sqrt{x}(\log x)^{1/3}$, so the left-hand side (LHS) of Eq. (32) becomes

$$\int_1^A \frac{1}{x(\log x)^{2/3}} = 3(\log A)^{1/3} + c \tag{33}$$

where c is uniformly bounded and independent of A. Next, focusing on the RHS of Eq. (32), we get

$$\begin{aligned}\frac{1}{A^{3/2}}\int_1^A \sqrt{x}(\log x)^{1/3}\,dx &= (\log A)^{1/3} - \frac{1}{A^{3/2}}\int_1^A x^{3/2}\frac{1}{3x(\log x)^{2/3}}\,dx\\ &= (\log A)^{1/3} - \frac{1}{A^{3/2}}\underbrace{\int_1^A \sqrt{x}\frac{1}{3(\log x)^{2/3}}\,dx}_{=o\left(\int_1^A \sqrt{x}\cdot(\log x)^{1/3}\,dx\right)}\\ &= \Theta((\log A)^{1/3}) \end{aligned}\tag{34}$$

Our uniqueness claim follows by noting that the LHS of Eq. (32) is decreasing in $f(x)$ (in orders of magnitude of the upper bound of the integral) while the right-hand side (RHS) is increasing in $f(x)$.

Case 2: $f(k) = o(g(k))$. For the second case, assume that $f(k) = o(g(k))$. This implies for the matching cost that[19]

$$\int_1^A \frac{1}{f(\tau)^2}\,d\tau = \omega\left(\int_1^A \frac{1}{g(\tau)^2}\,d\tau\right)\tag{35}$$

The integral on the RHS of Eq. (35) can then be bounded from below as follows

$$\int_1^A \frac{1}{g(\tau)^2}\,d\tau = \int_1^A \frac{1}{4\tau\log\log\tau}\,d\tau = \omega\left(\int_1^A \frac{1}{4\tau(\log\tau)^{2/3}}\,d\tau\right)\tag{36}$$

For the integral on the RHS of Eq. (36) we have

$$\int_1^A \frac{1}{4\tau(\log\tau)^{2/3}}d\tau = \Theta\big((\log A)^{1/3}\big),\tag{37}$$

Hence, combining these last approximations, we finally get

$$\int_1^A \frac{1}{f(\tau)^2}\,d\tau = \omega\big((\log A)^{1/3}\big).\tag{38}$$

Thus any solution satisfying $f(k) = o(g(k))$ (Case 2) has expected matching cost that is $\omega(1)$ relative to the optimal solution. This completes the proof that $f(k) = \Theta(\sqrt{x}(\log x)^{1/3})$ is the unique optimal clearing schedules for the balanced social planner. □

[19]Formally, for $\tau \geq e$, the integrand is not well-defined, but the Cauchy principal value of the integral remains finite, and this is the value we are using for $\tau \leq e$. This issue could be side-stepped by shifting the lower limit of the integral to a higher value, but we do not do so in order to simplify the presentation.

Note that, up to logarithmic factors, the balanced clearing schedule is close to the clearing schedule $\texttt{CS}_{\gamma=1/2}$ which signifies a first-order phase transition for the expected matching ratio. As discussed earlier, $\texttt{CS}_{\gamma=1/2}$ only signifies a second-order phase transition for the expected waiting ratio, thus explaining the gap between $\texttt{CS}_{\texttt{balanced}}$ and $\texttt{CS}_{\gamma=1/2}$. In practice however, $\texttt{CS}_{\gamma=1/2}$ seems to be a reasonable approximation for a balanced social planner.

6 Generalization

So far, we focused our attention on dynamic clearing games with exponentially distributed match costs. We shall show below that the developed techniques can also be used to study a more abstract model. Rather than modeling match costs directly by defining the distribution of each potential match cost, w_{ij}, we take a macroscopic viewpoint and posit that the cost of matching a couple depends on the number of clients and providers currently in the market. This cost may be the expected cost of matching the cheapest couple or a cost associated to market making more generally. In practice, this cost can be learned from past data on clearing events. We shall thus focus our generalization on identifying breaking points and their respective consequences for different cost regimes.

Write $M_{\mathcal{C}}(\tau) = N_{\mathcal{C}}(\tau) - A(\tau)$ for the number of clients in the market at time τ and $M_{\mathcal{P}}(\tau) = N_{\mathcal{P}}(\tau) - A(\tau)$ for the number of providers respectively. Then, the expected cost can be written w.l.o.g. as

$$g(M_{\mathcal{C}}, M_{\mathcal{P}}) \tag{39}$$

where $g\colon \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ is a non-increasing function (in either argument).[20] Intuitively, g determines how the expected minimum cost of matching decreases as more players coexist in the market. For example, if $g(M_{\mathcal{C}}, M_{\mathcal{P}}) = \frac{1}{M_{\mathcal{C}} \cdot M_{\mathcal{P}}}$, we revert to the previous analysis resulting from exponentially distributed match costs. This is the case since the expected minimum of $M_{\mathcal{C}} \cdot M_{\mathcal{P}}$ independent $\exp(1)$ random variables is equal to $\frac{1}{M_{\mathcal{C}} \cdot M_{\mathcal{P}}}$.

In view of this, it stands to reason that the asymptotic behavior of the market will be captured by the rate at which the expected minimum matching cost $g(M_{\mathcal{C}}, M_{\mathcal{P}})$ vanishes as a function of $M_{\mathcal{C}}, M_{\mathcal{P}} \to \infty$. Theorem 6 below makes this intuition

[20] We define the function g on the real numbers, but note that it is only the values on $\mathbb{N} \times \mathbb{N}$ which enter the analysis of clearing schedules.

precise and identifies a specific threshold beyond which it *is* possible to get a 'free lunch'. We restrict our analysis to the case $\delta > 1$ to guarantee that the patient clearing schedule has finite expected matching cost, i.e., $\text{cost}_{\texttt{patient}} < \infty$.

Theorem 6. *Suppose that the expected minimum matching cost decays as $g(\Theta(x), \Theta(x)) = \Theta(1/x^{\delta})$ for some $\delta > 1$. Then:*

(i) *For $1 < \delta \leq 2$ there is no 'free lunch'. In particular, the critical rate clearing schedule, that is, the clearing schedule with expected matching ratio $\Theta(\log(A))$, is given by $\texttt{CS}_{\gamma=1/\delta}$.*

(ii) *For $\delta > 2$, 'free lunch' exists. In particular, the clearing schedules $\texttt{CS}_{\gamma}$ with $\gamma \in (\frac{1}{\delta}, \frac{1}{2}]$ guarantee that the expected matching and waiting ratios are both finite.*

Proof of Theorem 6. We first consider the upper bound. Given $g(\Theta(x), \Theta(x)) = \Theta\left(\frac{1}{x^{\delta}}\right)$ and since g is increasing in both arguments we have:

$$\mathbb{E}\left[\sum_{k=1}^{A} g(M_{\mathcal{C}}, M_{\mathcal{P}}) \,\middle|\, \min\{M_{\mathcal{C}}, M_{\mathcal{P}}\} = \lfloor k^{1/\delta} \rfloor\right] \leq \sum_{k=1}^{A} g(\lfloor k^{1/\delta} \rfloor, \lfloor k^{1/\delta} \rfloor)$$
$$= \Theta\left(\sum_{k=1}^{A} \frac{1}{k}\right) = \Theta(\log A) \qquad (40)$$

where the last inequality follows from the bounds for the harmonic series. Thus, given the optimal clearing schedule has finite matching cost, the expected matching ratio is smaller than $O(\log A + 1)$.

For the lower bound note that $\mathbb{E}[t(k, k^{1/\delta})] < 10k$ by similar arguments as in Eq. (58) and by recalling that $\delta > 1$. Thus, combining Jensen's inequality ($\frac{1}{x}$ is convex) with Markov's inequality, Lemma 9, and recalling that $\delta \leq 2$ we get:

$$
\begin{aligned}
&\sum_{k=1}^{A} \mathbb{E}\big[g(M_{\mathcal{C}}, M_{\mathcal{P}}) \,\big|\, \min\{M_{\mathcal{C}}, M_{\mathcal{P}}\} = \lfloor k^{1/\delta} \rfloor\big] \\
&= \sum_{k=1}^{A} \mathbb{E}\big[g(M_{\mathcal{C}}, M_{\mathcal{P}}) \,\big|\, \min\{M_{\mathcal{C}}, M_{\mathcal{P}}\} = \lfloor k^{1/\delta} \rfloor,\ t(k, k^{1/\delta}) < 20k\big] \cdot \mathbb{P}\left[t(k, k^{1/\delta}) < 20k\right] \\
&+ \sum_{k=1}^{A} \mathbb{E}\big[g(M_{\mathcal{C}}, M_{\mathcal{P}}) \,\big|\, \min\{M_{\mathcal{C}}, M_{\mathcal{P}}\} = \lfloor k^{1/\delta} \rfloor,\ t(k, k^{1/\delta}) > 20k\big] \cdot \mathbb{P}[t(k, k^{1/\delta}) > 20k] \\
&\geq \sum_{k=1}^{A} \mathbb{E}\big[g(M_{\mathcal{C}}, M_{\mathcal{P}}) \,\big|\, \min\{M_{\mathcal{C}}, M_{\mathcal{P}}\} = \lfloor k^{1/\delta} \rfloor,\ t(k, k^{1/\delta}) < 20k\big] \cdot \frac{1}{2} \\
&= \Theta\left(\sum_{k=1}^{A} g(\lfloor k^{1/\delta} \rfloor, \lfloor k^{1/\delta} \rfloor)\right) = \Theta\left(\sum_{k=1}^{A} \frac{1}{k}\right) = \Theta(\log A) \qquad (41)
\end{aligned}
$$

where we used that $M_{\mathcal{C}} = \Theta(M_{\mathcal{P}})$ given $|S_t| = |N_{\mathcal{C}}(t) - N_{\mathcal{P}}(t)| = |M_{\mathcal{C}}(t) - M_{\mathcal{P}}(t)| = O(\sqrt{20k})$ since we are in the case $t(k, k^{1/\delta}) < 20k$ and by the assumption that for all λ, μ we have $g(\lambda \cdot x, \mu \cdot x) = \Theta(x^{\delta})$. Thus, given the optimal clearing schedule has finite matching cost, the expected matching ratio is $\Omega(\log A)$, concluding the proof together with the upper bound.

To summarize, for $\delta \in (1, 2]$ the critical rate clearing schedule is given by $\texttt{CS}_{\gamma=1/\delta}$. Thus the critical rate clearing schedules are given by $\texttt{CS}_{\gamma}$ with $\gamma \in (0, 1/2]$ and by Theorem 4, the expected waiting ratio for these schedules is not finite. Note that the case $\delta = 2$ is simply Theorem 2. We conclude that there is no 'free lunch'.

(2) By Theorem 4 the expected waiting ratio is finite for all clearing schedules $\texttt{CS}_{\gamma}$ with $\gamma \leq \frac{1}{2}$.

We upper bound the expected matching ratio for the clearing schedule $\texttt{CS}_{\gamma}$ for $\gamma > \frac{1}{\delta}$: For the upper bound we have with $g(\Theta(x), \Theta(x)) = \Theta\left(\frac{1}{x^{\delta}}\right)$:

$$
\begin{aligned}
\mathbb{E}\left[\sum_{k=1}^{A} g(M_{\mathcal{C}}, M_{\mathcal{P}}) \,\middle|\, \min\{M_{\mathcal{C}}, M_{\mathcal{P}}\} = \lfloor k^{1/\delta} \rfloor\right] &\leq \sum_{k=1}^{A} g(\lfloor k^{1/\delta} \rfloor, \lfloor k^{1/\delta} \rfloor) \\
&= \Theta\left(\sum_{k=1}^{A} \frac{1}{k^{\gamma \cdot \delta}}\right) = \Theta(1) \qquad (42)
\end{aligned}
$$

where the last identity holds since $\gamma > \frac{1}{\delta}$.

Thus, for given δ the clearing schedules $\texttt{CS}_{\gamma}$ with $\gamma \in (\frac{1}{\delta}, \frac{1}{2}]$ guarantee that the expected matching ratio and waiting ratio are both finite, i.e., free lunch. □

Theorem 6(i) extends our previous analysis by showing how the critical rate clearing schedule moves dependent on δ. In fact Theorem 6(ii) shows that the conclusion of Theorem 2 does not hold for the regime $\delta > 2$, that is, there exists a free lunch and in particular it is achieved by the clearing schedules $\mathtt{CS}_\gamma$ with $\gamma \in (\frac{1}{\delta}, \frac{1}{2}]$. This is because for quickly decaying matching costs it becomes easier to choose a 'good' schedule, and thus the 'window of opportunity' is increasing in the derivative of g. One can build intuition for this result by reasoning about market settings that differ in terms of match cost variability: in markets where match costs are generally rather similar, thickening the market by waiting will only lead to a meaningful positive effect in terms of expected match cost reduction when waiting for a long time. By contrast, when match costs vary substantially, match costs reduce in expectation with much less delay, thus making it more likely for a mechanism designer to get a free lunch. Importantly, the clearing schedule $\mathtt{CS}_{\mathtt{greedy}}$ is never optimal when dealing with many types, independent of the match cost distribution at hand.

7 Discussion

In this paper, we studied the *dynamic clearing game*, where heterogeneous clients and providers arrive uncoordinatedly in order to be matched. We studied the trade-off a social planner is facing between two competing objectives: a) to reduce players' *waiting time* before getting matched; and b) to form efficient pairs in order to reduce *matching cost*.

Our analysis of the dynamic clearing game reveals that a multi-objective social planner often faces a substantial trade-off. Starting with the micro-founded model for match costs we showed that there exists no free lunch, that is, there is no clearing schedule that is approximately optimal in terms of both waiting time and matching cost. We identified a unique breaking point where a stark reduction in matching cost compared to a stark increase in waiting cost occurs. In line with recent works by Ashlagi et al. (2017), Ashlagi et al. (2018) and many others, we focused on a concrete class of social welfare functions that weigh costs from waiting versus matching on a comparable scale and identify the optimal clearing schedule, namely, the clearing schedule that matches the k-th couple when $\Theta(\sqrt{k}(\log k)^{1/3})$ players are on the short side of the market.

Generalizing the model, we abstract away from modeling match costs directly and take a macroscopic viewpoint. Positing that the cost of matching a couple depends on the number of clients and providers currently in the market we identify two regimes. One, where no free lunch continues to hold, the other, where there is a window of opportunity to be optimal along both dimensions, that is free lunch.

There are multiple directions in which our analysis could be extended. Perhaps the most evident avenue for future research is to model market participation behavior game-theoretically, which would lead to new strategic considerations and probably induce other matchings (see, e.g., Baccara et al. 2018). This analysis could be pursued in more applied contexts, for instance relating to our motivating example of a labor market with a central employment bureau, where waiting costs could be interpreted as benefits payable by the bureau. An unemployed worker might forgo some of these benefits by (repeatedly) rejecting matches. This is the case because longer waiting, even though borne out of strategic behavior, may improve the match quality (reducing matching cost).

A second route for further investigation is to enlarge the options of the social planner in terms of clearing schedules. For one, the social planner could be learning from market observations about the distribution of match costs, which incidentally we may also allow to follow other, more general classes of distributions. This would allow the social planner to formulate more sophisticated clearing schedules that incorporate match costs between players that are currently in the market. In particular, if the social planner learns that a given agent may be 'hard to match', then it might be sensible to match that agent directly and not incur further waiting cost. Furthermore, the social planner may want to match more than one couple at a time.

The study of dynamic market institutions is clearly fascinating, with tremendous scope for progress in (old and new) applications, where research has only just started. Our contribution has been to go beyond binary match values, and to identify breaking points under incomplete information. We hope that our framework is able to provide fertile ground for further research, both theoretical and applied to real-world market contexts, in particular as regards thinking about whether the kinds of breaking points we describe are relevant in the optimal design of such markets.

Appendix

A Proof of Theorem 3

Before turning to the proof we introduce the following definition and lemmas.

Definition 7. Let $X_1, X_2, \dots$ be iid random variables with $\mathbb{P}[X_i = 1] = \mathbb{P}[X_i = -1] = \frac{1}{2}$.

- Let $S_k = \sum_{i=1}^{k} X_i$.
- Let $t(k, C)$ be the stopping time for the event that for the k-th time at least C clients and C providers are in the market, assuming that every time this is the case one client and one provider are removed.

Lemma 8. *Let $w_{ij} \sim \exp(\lambda_j)$ for $j = 1, 2, \dots, N$, be a family of independent exponentially distributed random variables. Then*

$$\min\{w_{i1}, w_{i2}, \dots, w_{iN}\} \sim \exp\left(\sum_{j=1}^{N} \lambda_j\right). \tag{43}$$

In particular, if for all j, $\lambda_j = 1$, then $\mathbb{E}[\min_j w_{ij}] = \frac{1}{N}$.

Proof. This proof is standard but we repeat it for the sake of completeness. The random variable w_{ij} has cumulative distribution function

$$F_{w_{ij}} = \mathbb{P}(w_{ij} \leq x) = 1 - e^{-\lambda_j x} \quad \text{for all } x > 0 \text{ and all } j = 1, 2, \dots, N. \tag{44}$$

Now, define the random variable $Y = \min\{w_{i1}, w_{i2}, \dots, w_{iN}\}$. Then, the cumulative distribution function of Y is

$$\begin{aligned}
F_Y(y) &= \mathbb{P}(Y \leq y) \\
&= 1 - \mathbb{P}(Y \geq y) \\
&= 1 - \mathbb{P}(\min\{w_{i1}, w_{i2}, \dots, w_{iN}\} \geq y) \\
&= 1 - \mathbb{P}(w_{i1} \geq y) \cdot \mathbb{P}(w_{i2} \geq y) \cdot \ldots \cdot \mathbb{P}(w_{iN} \geq y) \\
&= 1 - e^{-\lambda_1 y} \cdot e^{-\lambda_2 y} \cdot \ldots \cdot e^{-\lambda_N y} \\
&= 1 - e^{-\sum_{j=1}^{N} \lambda_j y} \qquad y > 0
\end{aligned} \tag{45}$$

The latter cumulative distribution function is that of an exponential variable with parameter $\sum_{j=1}^{N} \lambda_j$. □

Lemma 9. *For S_k defined as above we have:*[21]

$$0.67 \cdot \sqrt{k} \lessapprox \frac{2\pi}{e^2} \cdot \sqrt{\frac{2}{\pi}} \cdot \sqrt{k} \leq \mathbb{E}[|S_k|] \leq \frac{e}{\sqrt{\pi}} \cdot \sqrt{\frac{2}{\pi}} \cdot \sqrt{k} \lessapprox 1.23 \cdot \sqrt{k} \tag{46}$$

Proof. The starting point of our proof is an intermediate result in the proof of the limit of the expected absolute value of the 1-d random walk, which is detailed in Hizak and Logozar (2011, Equations 29a and 29b) and is based on combinatorial arguments via the binomial distribution:

$$\mathbb{E}[|S_k|] = \begin{cases} \frac{1}{2^{k-2}} \frac{k}{2} \binom{k-1}{k/2} = \frac{k}{2^k} \frac{k!}{[(k/2)!]^2} & \text{for } k \text{ even,} \\ \frac{1}{2^{k-1}} \frac{k+1}{2} \binom{k}{(k+1)/2} = \frac{k+1}{2^{k+1}} \frac{(k+1)!}{[((k+1)/2)!]^2} & \text{for } k \text{ odd.} \end{cases} \tag{47}$$

Since $\mathbb{E}[|S_{2k}|] = \mathbb{E}[|S_{2k-1}|]$ it suffices to analyze the case where k is even. To that end, we will use Stirling's formula to bound $k!$ from above and below as

$$\sqrt{2\pi} \cdot k^{k+1/2} \cdot e^{-k} \leq k! \leq e \cdot k^{k+1/2} \cdot e^{-k} \tag{48}$$

For k even, we may bound $|S_k|$ from above as:

$$\mathbb{E}[|S_k|] = \frac{k}{2^k} \frac{k!}{[(k/2)!]^2} \leq \frac{k}{2^k} \frac{e \cdot k^{k+\frac{1}{2}} \cdot e^{-k}}{2\pi \cdot (k/2)^{k+1} \cdot e^{-k}} = \frac{e}{\sqrt{2\pi}} \cdot \sqrt{\frac{2}{\pi}} \cdot \sqrt{k} \tag{49}$$

Next, we lower bound $|S_k|$ for k even:

$$\mathbb{E}[|S_k|] = \frac{k}{2^k} \frac{k!}{[(k/2)!]^2} \geq \frac{k}{2^k} \frac{\sqrt{2\pi} \cdot k^{k+1/2} \cdot e^{-k}}{e^2 \cdot (k/2)^{k+1} \cdot e^{-k}} = \frac{2\pi}{e^2} \cdot \sqrt{\frac{2}{\pi}} \cdot \sqrt{k} \tag{50}$$

This concludes the proof for k even. For k odd we have with the observation that $|S_k| = |S_{k+1}|$:

$$\frac{2\pi}{e^2} \cdot \sqrt{\frac{2}{\pi}} \cdot \sqrt{k} \leq \frac{2\pi}{e^2} \cdot \sqrt{\frac{2}{\pi}} \cdot \sqrt{2\lceil k/2 \rceil} \leq \mathbb{E}[|S_{k+1}|] = \mathbb{E}[|S_k|] \tag{51a}$$

and

$$\frac{e}{\sqrt{\pi}} \cdot \sqrt{\frac{2}{\pi}} \cdot \sqrt{k} \geq \frac{e}{\sqrt{\pi}} \cdot \sqrt{\frac{2}{\pi}} \cdot \frac{1}{\sqrt{2}} \cdot \sqrt{2\lceil k/2 \rceil} \geq \mathbb{E}[|S_{k+1}|] = \mathbb{E}[|S_k|] \tag{51b}$$

[21] Note that $\lim_{k\to\infty} \mathbb{E}[|S_k|] = \sqrt{\frac{2}{\pi}} \cdot \sqrt{k}$ (Peters, 1856).

$\square$

For the sake of limiting notation the proposition and proof are stated for the clearing schedules with $f(k) = \lceil k^\gamma \rceil$ rather than for $\Theta(f)$. Adding constant upper and lower bounds is straightforward and thus omitted. Recall that $t(k, f(k))$ is the stopping time for the event that for the k-th time at least $f(k)$ clients and $f(k)$ providers are in the market, assuming that every time this is the case one client and one provider are removed.

Proof of Theorem 3. Throughout the proof we shall simplify notation by omitting the fact that some of the matching schedules are defined via the ceiling of functions mapping to $\mathbb{R}^+$ (e.g., $\lceil k^\gamma \rceil$). The results are not changed by the omission since match costs are never underestimated and overestimated by very little. Further, while they are stated together in the proposition, we study the clearing schedules $\texttt{CS}\gamma = 0$ and $\texttt{CS}_{0<\gamma<1/2}$ separately since they require different arguments.

First come, first served ($\texttt{CS}_{\texttt{FCFS}}$) In FCFS the cost of each match is the expectation of a single match cost, that is λ. After A matches have occurred, the expected incurred cost is λA. Thus, given the patient clearing schedule has cost bounded above by $\frac{\pi^2}{6\underline{\lambda}}$, the expected matching ratio is equal to $\frac{\lambda \cdot A}{\pi^2/(6\underline{\lambda})}$.

Before stating the proofs for the other results recall from the proof of Proposition 1, that the exponential distribution is closed under scaling. We shall thus simplify notation and assume that for all i, j $w_{ij} \sim \exp(1)$. Note that, for lower bounds the scaling factor $\frac{1}{\lambda}$ needs to be applied and for upper bounds the scaling $\frac{1}{\underline{\lambda}}$ needs to be applied. But note that those scaling factors are constant with respect to τ (and thus A) and therefore do not influence the orders of the limiting results.

Greedy matching ($\texttt{CS}_{\texttt{greedy}}$) The k-th match happens when the minimum of the number of clients and providers who already arrived to the market is k, that is, at time $t(k, 1)$. The expected weight of the k-th match depends on the number of players currently present on the long side of the market (since on the short side there is only one agent). This random variable is given by $|S_{t(k,1)}| + 1$. By Lemma 8 the expected weight thus is $\mathbb{E}[\frac{1}{|S_{t(k,1)}|+1}]$. The first A matches thus have an expected cost of

$$\mathbb{E}[\sum_{k=1}^{A} \frac{1}{|S_{t(k,1)}| + 1}]. \tag{52}$$

Given that we study fixed A (the number of matches that) we have:[22]

$$\mathbb{E}\left[\sum_{k=1}^{A}\frac{1}{|S_{t(k,1)}|+1}\right]=\sum_{k=1}^{A}\mathbb{E}[\frac{1}{|S_{t(k,1)}|+1}] \tag{53}$$

Next, by Jensen's inequality ($\frac{1}{x}$ is convex) we have:

$$\sum_{k=1}^{A}\mathbb{E}[\frac{1}{|S_{t(k,1)}|+1}] > \sum_{k=1}^{A}\frac{1}{\mathbb{E}[|S_{t(k,1)}|+1]} = \sum_{k=1}^{A}\frac{1}{\mathbb{E}[|S_{t(k,1)}|]+1} \tag{54}$$

We shall now approximate $\mathbb{E}[t(k,1)]$. By Lemma 9 we have $\mathbb{E}[S_t] < 1.23\sqrt{t}$. Thus the short side of the market has $\frac{t-1.23\sqrt{t}}{2}$ agents. Setting $k=\frac{t-1.23\sqrt{t}}{2}$ and solving the quadratic equation we find the crude upper bound for the expectation:[23]

$$\mathbb{E}[t(k,1)] = (\frac{1.23+\sqrt{1.23^2+8k}}{2})^2 < \frac{3}{4}+2k+2\sqrt{k} < 5k \tag{55}$$

Returning to Eq. (54) we have with Lemma 9:

$$\begin{aligned}\sum_{k=1}^{A}\frac{1}{\mathbb{E}[|S_{t(k,1)}|]+1} &> \sum_{k=1}^{A}\frac{1}{\mathbb{E}[|S_{5k}|]+1} > \frac{1}{1.23}\sum_{k=1}^{A}\frac{1}{\sqrt{5k}+1} && (56)\\ &> \frac{1}{1.23}A\cdot\frac{1}{\sqrt{5A}+1} > A\cdot\frac{1}{5\sqrt{A}} = \frac{\sqrt{A}}{5} && (57)\end{aligned}$$

Thus, given the optimal schedule has $\text{cost}_{\texttt{patient}}(A) \leq \frac{\pi^2}{6\underline{\lambda}}$, the expected matching ratio is lower bounded by $\frac{\sqrt{A}}{5\pi^2/(6\underline{\lambda})}$.

Subcritical matching ($\texttt{CS}_{\gamma=0}$) We shall fix the clearing schedule such that it matches a couple every time some fixed $C \in \mathbb{N}$ players are on the short side of the market ($N - A = C$) and note that it belongs to the family of clearing schedules $\texttt{CS}_{\gamma=0}$.

Next, note that $t(k,C) = t(1,C) + t(k-1,1)$, since we assume that every time

[22] Note that t (the total number of client and providers who have arrived to the market) depends on A (and vice versa). Therefore, Wald (1944)'s equation does not apply and thus the route of inquiry to study the matching cost at some continuous time τ does not work since we could not interchange summation and expectation.

[23] We solve $k=\frac{t-1.23\sqrt{t}}{2}$. Setting $t=u^2$ and rearranging we solve quadratic equation $u^2-u-2k \overset{!}{=} 0$. The solutions are:

$$u_{1,2}=\frac{1.23\pm\sqrt{1.23^2+8k}}{2}$$

Given the variable transformation the positive solution is selected.

at least C clients and C providers are in the market exactly one client and one provider match and thus leave the market. Similarly to the proof of Theorem 3($\texttt{CS}_{\texttt{greedy}}$) we can bound $\mathbb{E}[t(1,C)]$ from above by noting that it is equal to $\mathbb{E}[t(C,1)]$. Thus

$$\mathbb{E}[t(k,C)] = \mathbb{E}[t(C,1)] + \mathbb{E}[t(k-1,1)] < 5C + 5k \tag{58}$$

With the latter and, as above by Jensen's inequality ($\frac{1}{x}$ is convex) and Lemma 9 we have for the expected matching cost:

$$\begin{aligned}
\mathbb{E}[\sum_{k=1}^{A} \frac{1}{(C+|S_{t(k,C)}|)C}] &\geq \sum_{k=1}^{A} \frac{1}{(C+\mathbb{E}[|S_{t(k,C)}|])C} \geq \sum_{k=1}^{A} \frac{1}{(C+\mathbb{E}[|S_{5C+5k}|])C} \\
&\geq \frac{1}{1.23} \sum_{k=1}^{A} \frac{1}{(C+\sqrt{5C+5k})C} \geq \frac{A}{1.23} \cdot \frac{1}{(C+\sqrt{5C+5A})C} \\
&= \frac{1}{1.23 \cdot C} \cdot \frac{A-C}{\sqrt{5C+5A}+C} = \Omega(\sqrt{A})
\end{aligned} \tag{59}$$

Thus, given the optimal clearing schedule has finite cost the expected matching ratio is $\alpha(A) = \Omega(\sqrt{A})$.

The second part of the assertion follows by observing:

$$\mathbb{E}[\sum_{k=1}^{A} \frac{1}{(C+|S_{t(k,C)}|)C}] < \frac{1}{C}\mathbb{E}[\sum_{k=1}^{A} \frac{1}{1+|S_{t(k,1)}|}] \tag{60}$$

Subcritical matching ($\texttt{CS}_{0<\gamma<1/2}$) For the upper bound, by Lemma 8, we have:

$$\begin{aligned}
\mathbb{E}[\sum_{k=1}^{A} \frac{1}{(k^\gamma+|S_{t(k,\sqrt{k})}|)k^\gamma}] &\leq \sum_{k=1}^{A} \frac{1}{k^{2\gamma}} = 1 + \sum_{k=2}^{A} \frac{1}{k^{2\gamma}} \leq 1 + \int_{x=1}^{A} \frac{1}{x^{2\gamma}}\, dx \\
&= 1 + \left[\frac{1}{1-2\gamma} x^{1-2\gamma}\right]_{x=1}^{A} \leq 1 + \frac{1}{1-2\gamma} A^{1-2\gamma}
\end{aligned} \tag{61}$$

Thus, given the optimal clearing schedule has finite cost, the expected matching ratio is $\alpha(A) = \mathcal{O}(A^{1-2\gamma})$ for $0 < \gamma < \frac{1}{2}$.

For the lower bound, note that $t(k, k^\gamma) < 10k$ for $\gamma < 1$ by similar arguments as in Eq. (58). Further note that $\mathbb{E}[|S_t|]$ is strictly increasing in t. Thus, with Jensen's inequality ($\frac{1}{x}$ is convex):

$$\begin{aligned}\sum_{k=1}^{A}\mathbb{E}[\frac{1}{(k^\gamma+|S_{t(k,k^\gamma)}|)k^\gamma}] &> \sum_{k=1}^{A}\frac{1}{(k^\gamma+\mathbb{E}[|S_{10k}|])k^\gamma} &&> \frac{1}{1.23}\sum_{k=1}^{A}\frac{1}{(k^\gamma+\sqrt{10k})k^\gamma}\\ &> \frac{1}{1.23(\sqrt{10}+1)}\sum_{k=1}^{A}\frac{1}{k^{\frac{1}{2}+\gamma}} &&> \frac{1}{6}\int_{x=1}^{A}\frac{1}{x^{\frac{1}{2}+\gamma}}dx\\ &> \frac{1}{6}[\frac{1}{\frac{1}{2}-\gamma}x^{\frac{1}{2}-\gamma}]_{x=1}^{A} &&= \Omega(A^{\frac{1}{2}-\gamma}) \end{aligned} \tag{62}$$

Critical matching ($\texttt{CS}_{\gamma=1/2}$) For the upper bound, by Lemma 8, we have:

$$\mathbb{E}[\sum_{k=1}^{A}\frac{1}{(\sqrt{k}+|S_{t(k,\sqrt{k})}|)\sqrt{k}}] < \sum_{k=1}^{A}\frac{1}{k} \leq \log A+1 \tag{63}$$

where the last inequality follows from the bounds for the harmonic series. Thus, given the optimal clearing schedule has finite matching cost (lower bounded by $\frac{\log(2)}{\bar{\lambda}}$), the expected matching ratio is smaller than $\frac{\frac{1}{\underline{\lambda}}(\log A+1)}{\log(2)/\bar{\lambda}}$.

For the lower bound note that $\mathbb{E}[t(k,\sqrt{k})] < 10k$ by similar arguments as in Eq. (58). Further note that S_t is strictly increasing in t. Thus, with Jensen's inequality ($\frac{1}{x}$ is convex) and Lemma 9:

$$\begin{aligned}\sum_{k=1}^{A}\mathbb{E}[\frac{1}{(\sqrt{k}+|S_{t(k,\sqrt{k})}|)\sqrt{k}}] &> \sum_{k=1}^{A}\frac{1}{(\sqrt{k}+\mathbb{E}[|S_{10k}|])\sqrt{k}} > \frac{1}{1.23}\sum_{k=1}^{A}\frac{1}{(\sqrt{k}+\sqrt{10k})\sqrt{k}}\\ &\geq \frac{1}{6}\sum_{k=1}^{A}\frac{1}{k} > \frac{1}{6}\log A \end{aligned} \tag{64}$$

Thus, given the optimal clearing schedule has $\text{cost}_{\texttt{patient}}(A) \leq \frac{\pi^2}{6\underline{\lambda}}$, the expected matching ratio is bounded below by $\frac{\frac{1}{6}\log A}{\pi^2/(6\underline{\lambda})} = \frac{\underline{\lambda}}{\pi^2}\cdot\log A$.

Supercritical matching ($\texttt{CS}_{1/2<\gamma\leq 1}$) As above, by Jensen's inequality (since $1/x$ is convex) and Lemma 8 we have:

$$\mathbb{E}[\sum_{k=1}^{A}\frac{1}{(k^\gamma+|S_{t(k,k^\gamma)}|)k^\gamma}] < \sum_{k=1}^{A}\mathbb{E}[\frac{1}{k^\gamma k^\gamma}] = \sum_{k=1}^{A}\frac{1}{k^{2\gamma}} \to \zeta(2\gamma) \tag{65}$$

where ζ is the Riemann zeta function and is known to converge for $\gamma > \frac{1}{2}$. Given that we are considering a sum with positive summands convergence is from below.

Thus, given the optimal clearing schedule has finitematching cost ($\mathrm{cost}_{\texttt{patient}}(A) \geq \frac{\log(2)}{\overline{\lambda}}$), the expected matching ratio is bounded from above by $(\overline{\lambda}/\underline{\lambda}) \cdot \zeta(2\gamma)/\log 2$ for $\frac{1}{2} < \gamma \leq 1$. $\square$

B Proof of Theorem 4

Proof of Theorem 4. First note that in order to compare different clearing schedules we are interested in the additional waiting time incurred until some number A of couples have been matched. Thus, we consider the waiting time until T for the greedy schedule (the benchmark) and for other schedules the waiting time until $\hat{T}$ where $\hat{T}$ is the expected time until under the given schedule the same number of couples have been matched as in the greedy schedule until time T.

As in the Proof of Theorem 3 we shall simplify notation by omitting the fact that some of the matching schedules are defined via the ceiling function of functions mapping to $\mathbb{R}^+$ (e.g., $\lceil k^\gamma \rceil$). We invite the reader to convince her-or himself that the results are not altered through this simplification.

Let $\tau(k)$ be the moment the k-th couple is matched (given a particular clearing schedule). We proceed in a case-by-case basis below:

First come, first served ($\mathtt{CS}_{\mathtt{FCFS}}$) It suffices to note that this clearing schedules matches players at exactly the same moments as the greedy clearing schedules. The result then follows.

Subcritical and critical matching ($\mathtt{CS}_{0\leq\gamma\leq 1/2}$) We shall study the worst case such clearing schedule with respect to waiting time. We consider two different parts. In the first part we wait until at least T^γ clients and T^γ providers are in the market. The second part then proceeds in the same way as the greedy clearing schedule, keeping in mind that at all future times $\min\{N_{\mathcal{C}}, N_{\mathcal{P}}\}$ is exactly T^γ. The expected waiting time of the first schedule can be bounded above by the upper bound for the expected time until T^γ clients and T^γ providers are in the market, that is, $\mathbb{E}[\tau(5T^\gamma)] = 5T^\gamma$ (see Eq. (55) in the proof of Theorem 3) noting that we used the fact that the arrival of agents is governed by a Poisson clock of rate 1. Now, a crude upper bound for the waiting time of the first part of the process is found be assuming that all agents are in the market from the beginning ($\tau = 0$), yielding the upper bound $5T^\gamma \cdot 5T^\gamma$.

Note that, the first part of the process takes $\hat{T} - T$ time. For the remaining second part of the process the waiting cost is the cost of the greedy schedule ($\frac{2}{3}T^{3/2}$) plus the cost of the – in expectation – no more than $5T^\gamma$ agents on each side of the market to 'remain' for the subsequent periods. Thus the total waiting time is bounded above by:

$$5T^\gamma \cdot 5T^\gamma + \frac{2}{3}T^{3/2} + 5T^\gamma \cdot T = \Theta(T^{3/2}) \tag{66}$$

Thus $\beta(\hat{T}) = (3/2)\,\Theta(T^{3/2})/T^{3/2} = \Theta(1)$.

Supercritical matching ($\texttt{CS}_{1/2<\gamma\leq 1}$) We first construct a lower bound. Consider the alternative arrival process, where clients and providers alternatingly arrive to the market. Note that for any given clearing schedule this process incurs lower waiting time. For the clearing schedule we consider the waiting time of this alternative arrival process is precisely governed by the fact that the k-th match takes place when at least k^γ players are on the short side of the market. Further note that $\hat{T} \geq T$. Thus, the waiting time is lower bounded by using the approximation by the Wiener process (by arguments as in Proposition 1 and by observing that arrival is governed by a Poisson clock of rate 1):

$$\int_0^T 2\tau^\gamma d\tau = \frac{2}{1+\gamma}\tau^{1+\gamma}|_0^T = \Omega(T^{1+\gamma}) \tag{67}$$

For the upper bound, we construct a clearing schedule that constitutes an upper bound of the schedule under consideration. For fixed k, let $T = \tau(k)$ consider the following clearing schedule: First wait until there are at least k^γ clients and providers in the market, then proceed with the greedy schedule such that at any future point $\min\{clients, providers\}$ in the market is equal to k^γ. Note that this new schedule has the same total run time as the original schedule, that is, $\hat{T}$. Further it is evident that the waiting time occurred by the new schedule is greater than the waiting time of the original schedule. By arguments as for ($\texttt{CS}_{\gamma=0}$) and by the fact that arrival is governed by a Poisson clock of rate 1 we can upper bound the waiting time by:

$$5T^\gamma \cdot 5T^\gamma + \frac{2}{3}T^{3/2} + 5T^\gamma \cdot T = \Theta(T^{1+\gamma}) \tag{68}$$

since we assumed $\frac{1}{2} < \gamma \leq 1$.

The two bounds together show that the waiting time of the originally considered clearing schedule is $\Theta(T^{1+\gamma})$.Thus $\beta(\hat{T}) = \frac{\Theta(T^{1+\gamma})}{(2/3)\,T^{3/2}} = \Theta(T^{\gamma-\frac{1}{2}})$.

Patient matching ($\texttt{CS}_{\texttt{patient}}$) First note that for the patient schedule $\hat{T} = T$. The expected waiting time for the patient schedule until time T is given by

$$\mathbb{E}\left[\int_0^T N_{\mathcal{C}}(\tau) + N_{\mathcal{P}}(\tau)\, d\tau\right] = \int_0^T \mathbb{E}[N_{\mathcal{C}}(\tau) + N_{\mathcal{P}}(\tau)]\, d\tau \tag{69}$$

where the latter equality holds by Tonelli's theorem (by noting that $N_{\mathcal{C}}(\tau)+N_{\mathcal{P}}(\tau)$ is non-negative). The expectation is with respect to the number of clients and providers and with respect to the arrival times of the agents (governed by a Poisson clock). Again by Tonelli's theorem we can consider the case where the expectation with respect to the arrival times is taken first. Then by the fact that the arrival of agents is assumed to follow a Poisson clock of rate 1 we have:

$$\int_0^T \mathbb{E}[N_{\mathcal{C}}(\tau) + N_{\mathcal{P}}(\tau)]d\tau = \int_0^T \lfloor \tau \rfloor d\tau = \Theta(T^2) \tag{70}$$

Thus $\beta(\hat{T}) = \frac{\Theta(T^2)}{\frac{2}{3}T^{3/2}} = \Theta(\sqrt{T})$. □

C Proof of approximation in Proof of Theorem 5

Proof of omitted approximation in Proof of Theorem 5. We begin by approximating the two sums in Eq. (31), i.e.,

$$\sum_{k=1}^{A} \frac{1}{f(k)^2} = \Theta\left(\frac{1}{A^{3/2}} \sum_{k=1}^{A} f(k)\right) \tag{71}$$

Recalling that f is assumed non-decreasing for large k, the summand on the left-hand side is decreasing and

$$\int_0^A \frac{1}{f(x)^2}\, dx \geq \sum_{k=1}^{A} \frac{1}{f(k)^2} \geq \int_1^{A+1} \frac{dx}{f(x)^2} \tag{72}$$

Considering the meaning of $f(k)$ it is without loss of generality to define $f(x) = 1$ for $x \in [0, 1)$ since the summand $\frac{1}{f(k)^2}$ remains decreasing. Thus the absolute difference between the two bounds is bounded above by:

$$\left|\int_0^A \frac{1}{f(x)^2}\, dx - \int_1^{A+1} \frac{1}{f(x)^2}\, dx\right| = \left|\int_0^1 \frac{1}{f(x)^2}\, dx - \int_A^{A+1} \frac{1}{f(x)^2}\, dx\right| \leq 1 \tag{73}$$

It follows that

$$\sum_{k=1}^{A} \frac{1}{f(k)^2} = \Theta\left(\int_1^{A+1} \frac{1}{f(x)^2}\, dx\right) \tag{74}$$

Next consider the right-hand side of (31). The summand is increasing, so we get:

$$\frac{1}{A^{3/2}} \int_0^A f(x)\, dx \leq \frac{1}{A^{3/2}} \sum_{k=1}^{A} f(k) \leq \frac{1}{A^{3/2}} \int_1^{A+1} f(x)\, dx \tag{75}$$

Now note that $f(x) < x$ must hold. Thus the absolute difference between the two bounds is bounded above by:

$$\begin{aligned} \frac{1}{A^{3/2}}\left|\int_0^A f(x)\, dx - \int_1^{A+1} f(x)\, dx\right| &= \frac{1}{A^{3/2}}\left|\int_A^{A+1} f(x) dx - \int_0^1 f(x)\, dx\right| \\ &\leq \frac{A}{A^{3/2}} = \mathcal{O}(1). \end{aligned} \tag{76}$$

It follows that

$$\frac{1}{A^{3/2}} \sum_{k=1}^{A} f(k) = \Theta\left(\frac{1}{A^{3/2}} \int_1^{A+1} f(x)\, dx\right) \tag{77}$$

With above approximations it follows that Eq. (31) holds if and only if the following equation holds:

$$\int_1^A \frac{1}{f(x)^2}\, dx = \Theta\left(\frac{1}{A^{3/2}} \int_1^A f(x)\, dx\right) \tag{78}$$

as claimed. □

References

Aggarwal, G., G. Goel, C. Karande, and A. Mehta (2011), "Online vertex-weighted bipartite matching and single-bid budgeted allocations." In *Proceedings of the twenty-second annual ACM-SIAM symposium on discrete algorithms*, 1253–1264.

Akbarpour, M., S. Li, and S. Oveis Gharan (2017), "Thickness and information in dynamic matching markets." Working paper.

Aldous, D. J. (2001), "The $\zeta(2)$ limit in the random assignment problem." *Random Structures Algorithms*, 18, 381–418.

Anderson, R., I. Ashlagi, D. Gamarnik, and Y. Kanoria (2015), "A dynamic model of barter exchange." In *Proceedings of the twenty-sixth annual ACM-SIAM Symposium on Discrete Algorithms*, 1925–1933.

Ashlagi, I., Y. Azar, M. Charikar, A. Chiplunkar, O. Geri, H. Kaplan, R. Makhijani, Y. Wang, and R. Wattenhofer (2017), "Min-cost bipartite perfect matching with delays." *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2017)*, 81, 1–20.

Ashlagi, I., M. Burq, P. Jaillet, and V. H. Manshadi (2018), "On matching and thickness in heterogeneous dynamic markets."

Ashlagi, I., A. Nikzad, and P. Strack (2019), "Matching in dynamic imbalanced markets." Working paper.

Azar, Y., A. Chiplunkar, and H. Kaplan (2017), "Polylogarithmic bounds on the competitiveness of min-cost perfect matching with delays." In *Proceedings of the twenty-eight annual ACM-SIAM Symposium on Discrete Algorithms*, 1051–1061.

Baccara, M., S. Lee, and L. Yariv (2018), "Optimal dynamic matching." Working paper.

Bloch, F. and N. Houy (2012), "Optimal assignment of durable objects to successive agents." *Economic Theory*, 51, 13–33.

Buck, M. W., C. S. Chan, and D. P. Robbins (2002), "On the expected value of the minimum assignment." *Random Structures Algorithms*, 21, 33–58.

Budish, E., P. Cramton, and J. Shim (2015), "The high-frequency trading arms race: Frequent batch auctions as a market design response." *The Quarterly Journal of Economics*, 130, 1547–1621.

Dickhaut, J., S. Lin, D. Porter, and V. Smith (2012), "Commodity durability, trader specialization, and market performance." *PNAS*, 109, 1425–1430.

Dow, J. C. R. and L. A. Dicks-Mireaux (1958), "The excess demand for labour. A study of conditions in Great Britain, 1946-56." *Oxford Economic Papers*, 10, 1–33.

Edmonds, J. (1965), "Paths, trees, and flowers." *Canadian Journal of Mathematics*, 17, 49–467.

Egervary, E. (1931), "Matrixok kombinatorius tulajdonsagairol." *Matematikai es Fizikai Lapok*, 38, 16–27.

Emek, Y., S. Kutten, and R. Wattenhofer (2016), "Online matching: Haste makes waste!" In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, 333–344.

Emek, Y., Y. Shapiro, and Y. Wang (2019), "Minimum cost perfect matching with delays for two sources." *Theoretical Computer Science*, 754, 122–129.

Gale, D. and L. S. Shapley (1962), "College admissions and stability of marriage." *American Mathematical Monthly*, 69, 9–15.

Hardy, G. H. (1910), *Orders of infinity*. Cambridge University Press.

Hizak, J. and R. Logozar (2011), "A derivation of the mean absolute distance in one-dimensional random walk." *Tehnicki glasnik*, 5, 10–16.

Janson, S. (1999), "One, two and three times log n/n for paths in a complete graph with random weight." *Probability and Computing*, 8, 347–361.

Kac, M. (1947), "Random walk and the theory of Brownian motion." *The American Mathematical Monthly*, 54, 369–391.

Kanoria, Y., D. Saban, and J. Sethuraman (2018), "Convergence of the core in assignment markets." *Operations Research*, 66, 620–636.

Karp, R. M., U. V. Vazirani, and V. V. Vazirani (1990), "An optimal algoritm for on-line bipartite matching." In *Proceedings of the twenty-second annual ACM symposium on theory of computing*, 352–358.

Koenig, D. (1931), "Grafok es matrixok." *Matematikai es Fizikai Lapok*, 38, 116–119.

Krokhmal, P. A. and P. M. Pardalos (2009), "Random assignment problems." *European Journal of Operational Research*, 194, 1–17.

Kurino, M. (2014), "House allocation with overlapping generations." *American Economic Journal: Microeconomics*, 6, 258–289.

Leshno, J. (2012), "Dynamic matching in overloaded systems." Working paper.

Linusson, S. and J. Waestlund (2004), "A proof of Parisi's conjecture on the random assginment problem." *Probability Theory and Related Fields*, 128, 419–440.

Loertscher, S., E. V. Muir, and P. G. Taylor (2018), "Optimal market thickness and clearing." Working paper.

Lucas, R. and E. Prescott (1974), "Equilibrium search and unemployment." *Journal of Economic Theory*, 7, 188–209.

McCall, J. J. (1970), "Economics of information and job search." *The Quarterly Journal of Economics*, 84, 113–126.

Mehta, A. (2013), "Online matching and ad allocation." *Foundations and Trends in Theoretical Computer Science*, 8, 265–368.

Mezard, M. and G. Parisi (1987), "On the solution of the random link matching problems." *J. Physique*, 48, 1451–1459.

Mortensen, D. T. (1970), "A theory of wage and employment dynamics." *Microeconomic foundations of employment and inflation theory*, 167–211.

Myerson, R. B. and M. A. Satterthwaite (1983), "Efficient mechanisms for bilateral trading." *Journal of Economic Theory*, 29, 265–281.

Peters, C. A. F. (1856), "Uber die Bestimmung des wahrscheinlichen Fehlers einer Beobachtung aus den Abweichungen der Beobachtungen von ihrem arithmetischen Mittel." *Astronomische Nachrichten*, 44, 33–31.

Roth, A. E. (2012), "Marketplace institutions related to the timing of transactions: Reply to priest." *Journal of Labor Economics*, 30, 479–494.

Roth, A. E. and X. Xing (1994), "Jumping the gun: Imperfections and institutions related to the timing of market transactions." *American Economic Review*, 84, 992–1044.

Roth, A. E. and X. Xing (1997), "Turnaround time and bottlenecks in market clearing: Decentralized matching in the market for clinical psychologists." *Journal of Political Economy*, 105, 284–329.

Rustichini, A., M. A. Satterthwaite, and S. R. Williams (1994), "Convergence to efficiency in a simple market with incomplete information." *Econometrica*, 62, 1041–1063.

Shapley, L. S. and M. Shubik (1972), "The assignment game 1: The core." *International Journal of Game Theory*, 1, 111–130.

Shimer, R. (2007), "Mismatch." *American Economic Review*, 97, 1074–1101.

Smith, V. (1962), "An experimental study of competitive market behavior." *Journal of Political Economy*, 70, 111–137.

Smith, V. L., G. L. Suchanek, and A. W. Williams (1988), "Bubbles, crashes, and endogenous expectations in experimental spot asset markets." *Econometrica: Journal of the Econometric Society*, 1119–1151.

Uenver, M. U. (2010), "Dynamic kidney exchange." *The Review of Economic Studies*, 77, 372–414.

Waestlund, J. (2005), "A proof of a conjecture of buck, chan, and robbins on the expected value of the minimum assignment." *Random Structures Algorithms*, 26, 237–251.

Waestlund, J. (2009), "An easy proof of the $\zeta(2)$ limit in the random assignment game problem." *Electronic Communications in Probability*, 14, 261–269.

Wah, E., D. Hurd, and M. Wellman (2015), "Strategic market choice: Frequent call markets vs. continuous double auctions for fast and slow traders." *EAI Endorsed Transactions on Serious Games*, 3.

Wah, E., M. Wright, and M. P. Wellman (2017), "Welfare effects of market making in continuous double auctions." *Journal of Artificial Intelligence Research*, 59, 613–650.

Wald, A. (1944), "On cumulative sums of random variables." *The Annals of Mathematical Statistics*, 15, 283–296.

Walkup, D. W. (1979), "On the expected value of a random assignment problem." *SIAM Journal on Computing*, 8, 440–442.

Zenios, S. A. (2002), "Optimal control of a paired-kidney exchange program." *Management Science*, 48, 328–342.